Working Paper 94-30
Statistics and Econometrics Series 11
September 1994

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

SEMIPARAMETRIC TESTING OF NON-NESTED MODELS: AN APPLICATION TO ENGEL CURVES SPECIFICATION

Juan Mora*

Abstract

This paper proposes a specification test for non-nested semiparametrically specified competing models. The test-statistic is based on an artificial regression procedure. We derive the asymptotic distribution of the test-statistic under the null and alternative hypotheses and the finite-sample properties of the test are studied by means of simulations. The test is applied to discriminate between alternative Engel Curve specifications in share form, which relate total expenditure (X) with percentage of total expenditure spent on a specific good (Y). With data from the 1980 Spanish Family Expenditure Survey, one of the considered forms explains adequately the behaviour of households not contained on the tails of the distribution of X; when the whole data set is used both specifications are rejected possibly due to the fact that seemingly there is not mean dependence between Y and X for households with high total expenditure.

Key Words
Non-nested tests; Artificial nesting; Semiparametric regression; k-nearest neighbours; Engel Curves.

*Departamento de Estadística y Econometría, Universidad Carlos III de Madrid.
This paper is based on research funded by Spanish Dirección General de Investigación Científica y Técnica (DGCIYT), reference number PB92-0247. I am grateful to Miguel A. Delgado for his comments and suggestions.

1. INTRODUCTION

In recent years several procedures have been proposed for estimating the unknown parameter of the partially linear regression function

$$E[Y|X,Z] = \beta'X + g(Z) \quad a.s., \tag{1.1}$$

where (Y,X,Z) is an $\mathbb{R}\times\mathbb{R}^p\times\mathbb{R}^q$-valued observable random variable, β is an $\mathbb{R}^p$-valued unknown parameter vector and $g{:}\mathbb{R}^q \longrightarrow \mathbb{R}$ is an unknown real function. These procedures are discussed, among others, in Delgado and Mora (1994) (see references therein). However, there has been no work on the problem of model selection within the framework of equation (1.1). This paper deals with discrimination between non-nested partially linear regression models.

Pesaran and Deaton (1978) and Davidson and Mackinnon (1981) proposed procedures to test non-nested hypotheses. Pesaran and Deaton proposed a test-statistic based on an application of Cox's centred log-likelihood ratio criterion (Cox 1961, 1962). Davidson and Mackinnon proposed much simpler procedures which arise from artificial nesting (AN) of regression equations. These test-statistics based upon AN procedures, which are also straightforward applications of Cox's criterion (see White 1982 or Fisher 1983), have been extensively used in the econometric literature in recent years (see MacKinnon 1990). Our specification test for non-nested partially linear regression models is also based on the AN procedure.

We apply the proposed specification test to analise the validity of some forms of Engel Curves. Popular forms of Engel Curves are

$$E[Y|X] = \beta_0 + \beta_1(log\ X) + \beta_2(log\ X)^2,\ a.s., \tag{1.2}$$

where X is the total expenditure of a household and Y is the budget share spent on a certain good, and

$$E[Y|X] = \beta_o + \beta_1 X + \beta_2 X^{-1} \text{ a.s.} \tag{1.3}$$

Equation (1.2) generalises the Working-Leser (WL) form of Engel Curves (Working 1943 and Leser 1963) and has been used by Deaton (1981) and Deaton et al. (1989) among others. Equation (1.3) is the share form of Engel Curves deduced from the Quadratic Expenditure System (QES) of Pollack and Wales (1978, 1980). We test the validity of (1.2) and (1.3) taking into account other possible relevant variables Z (e.g: size of a household or age of reference person). If we assume that the relationship between Z and X is additive, then we can test the validity of the WL form and the QES form of Engel Curves using the specification test for non-nested partially linear regression models which we introduce in this paper.

In Section 2 we define the test statistic, prove its asymptotic properties and present some Monte Carlo results which show the finite-sample performance of the test in different sampling schemes. In Section 3 the proposed test is applied to analyse the validity of Engel Curves (1.2) and (1.3) using data from the 1980 Spanish Family Expenditure Survey. Proofs are confined to an appendix.

2. A SEMIPARAMETRIC SPECIFICATION TEST

The objective of this section is to propose a test statistic for choosing between alternative non-nested partially linear regression models. Suppose we have independent identically distributed observations $\{(Y_i,X_i,Z_i),\ 1\le i\le n\}$ from an $\mathbb{R}\times\mathbb{R}^p\times\mathbb{R}^q$-valued random variable (Y,X,Z) where $X = (X_1,X_2,X_3)$ takes values on $\mathbb{R}^k\times\mathbb{R}^l\times\mathbb{R}^m$ $(k+l+m = p)$. The researcher faces the competing hypotheses,

$$H_0: \qquad E[Y|X,Z] = X_1'\beta_1 + X_3'\beta_3 + g(Z) \quad \text{a.s.,} \tag{2.1}$$

$$H_1: \qquad E[Y|X,Z] = X_2'\beta_2 + X_3'\beta_3 + g(Z) \quad \text{a.s.,} \tag{2.2}$$

where $g(.)$ is an unknown function from $\mathbb{R}^q$ to $\mathbb{R}$ and β_1, β_2, β_3 are vectors of unknown parameters. In other words, the researcher has to decide between two alternative groups of variables in the linear part of a regression function in a situation where stacking all the independent variables and propose $E[Y|X,Z]=X'\beta + g(Z)$ is not sensible from an economic point of view. This is the case, for instance, in the specification of Engel Curves considered in Section 3, where X_1, X_2 and X_3 are income related variables and Z is a vector of regressors explaining personal characteristics.

In order to define a statistic for our test, first of all we must specify a procedure to estimate the coefficient β in the partially linear regression model (1.1). This model has been studied by many authors in recent years (see Delgado and Mora 1994 and references therein). Here we follow the approach of Robinson (1988) and Speckman (1988), who proposed feasible estimates of β based on nonparametric estimates of the unknown regression functions $E[Y|Z]\equiv m_Y$ and $E[X|Z]\equiv m_X$.

The idea behind the estimate they proposed may be explained as follows: equation (1.1) may be rewritten as

$$Y-m_Y = (X-m_X)'\beta + U, \tag{2.3}$$

where $E[U|X,Z] = 0$, *a.s.* If m_Y and m_X were known, (2.3) would be an ordinary regression model and, given a random sample $\{(Y_i,X_i,Z_i),\ 1\le i\le n\}$, the OLS procedure would give the root-n-consistent estimate[2]

$$\bar{\beta} \equiv \{\textstyle\sum_i (X_i-m_{Xi})(X_i-m_{Xi})'\}^{-1}\sum_i (X_i-m_{Xi})(Y_i-m_{Yi}).$$

In our model, we do not assume that the regression functions $m_Y(.)$ and $m_X(.)$ are of known functional form and, hence, $\bar{\beta}$ is infeasible. Feasible parameter estimates can be constructed from nonparametric estimates of the regression functions. Thus, we will consider

$$\hat{\beta} \equiv \{\textstyle\sum_i (X_i-\hat{X}_i)(X_i-\hat{X}_i)'I_i\}^{-1}\sum_i (X_i-\hat{X}_i)(Y_i-\hat{Y}_i)I_i, \tag{2.4}$$

where $\hat{X}_i$ and $\hat{Y}_i$ denote, respectively, nonparametric estimates of $m_{Xi}\equiv$

---

[2] Throughout this paper, all summations run from 1 to n unless otherwise specified. We also arbitrarily define $0/0$ to be 0, and the same convention applies whenever the inverse of a singular matrix appears.

$E[X_i|Z_i]$ and $m_{Yi} \equiv E[Y_i|Z_i]$ and I_i is a trimming function introduced for technical reasons.

The asymptotic properties of $\hat{\beta}$ have been studied, among others, by Robinson (1988), Speckman (1988) and Delgado and Mora (1994) under different sampling schemes. In the model analysed by Speckman (1988), Z is a fixed-design non-random variable, no trimming is required and any nonparametric estimate satisfying certain standard assumptions may be used. Robinson (1988) studies equation (1.1) when Z is an absolutely continuous random variable and errors are independent of regressors. He uses higher order kernels when q (dimension of Z) is greater than 1. The main result in Robinson (1988) is that, under certain regularity conditions,

$$n^{1/2}(\hat{\beta}-\beta) \xrightarrow{d} N(0,\sigma^2\Phi^{-1}), \tag{2.5}$$

where $\sigma^2 = Var(Y|X,Z)$ and $\Phi = E[(X-m_X)(X-m_X)']$. His theorem assumes independence between regressors and regression errors and, hence, it is not straightforwardly applicable to the heteroskedastic case. In Delgado and Mora (1994) equation (1.1) is considered first when Z contains only discrete random variables and then when Z contains both discrete and absolutely continuous random variables. In the former case independence between regressors and regression errors is not required. Hence, their main result can be easily generalised to the heteroskedastic partially linear regression model, in which case, under certain regularity conditions

$$n^{1/2}(\hat{\beta}-\beta) \xrightarrow{d} N(0,\Psi), \tag{2.6}$$

where $\Psi = E[(X-m_X)(X-m_X)']^{-1}E[\sigma^2(X,Z)(X-m_X)(X-m_X)']E[(X-m_X)(X-m_X)']^{-1}$ and $\sigma^2(X,Z) = Var(Y|X,Z)$.

2.1. Test-statistic

As mentioned above, the proposed test-statistic is based on an AN procedure. There are different ways to implement this procedure (see Davidson and Mackinnon 1981). Here, we artificially link the two competing hypotheses by means of the "composite hypothesis"

$$H_C:\ E[Y|X,Z] = (1-\delta)(X_1'\beta_1+X_3'\beta_3+g(Z)) + \delta(X_2'\beta_2+X_3'\beta_3+g(Z)),$$

In terms of H_C the two competing hypotheses (2.1) and (2.2) become

$$H_0:\quad \delta = 0, \tag{2.7}$$

$$H_1:\quad \delta = 1. \tag{2.8}$$

After a suitable reparametrization, the composite hypothesis can be rewritten as $E[Y|X,Z] = X_1'\gamma_1 + X_2'\beta_2\delta + X_3'\beta_3 + g(Z)$. Obviously, δ and β_2 are not identifiable in this equation. However, under H_1 it is possible to obtain a consistent estimate $\hat{\beta}_2$ of β_2 by using equation (2.4). We can then consider the artificial regression

$$Y = X_1'\gamma_1 + X_2'\hat{\beta}_2\delta + X_3'\beta_3 + g(Z) + U \tag{2.9}$$

and finally estimate the coefficients $(\gamma_1,\delta,\beta_3)$ of this partially linear model, using again equation (2.4). As suggested by (2.7) and (2.8), the t-ratio obtained for δ is the test-statistic.

Let us obtain the expression of the test statistic. Given a random sample $\{(Y_i,X_i,Z_i),\ 1\le i\le n\}$, first we obtain an estimate of β_2 from

$$\begin{pmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = \left\{ \sum_i \begin{pmatrix} \hat{\varepsilon}_{2i}\hat{\varepsilon}_{2i}' & \hat{\varepsilon}_{2i}\hat{\varepsilon}_{3i}' \\ \hat{\varepsilon}_{3i}\hat{\varepsilon}_{2i}' & \hat{\varepsilon}_{3i}\hat{\varepsilon}_{3i}' \end{pmatrix} I_i \right\}^{-1} \sum_i \begin{pmatrix} \hat{\varepsilon}_{2i}\hat{\varepsilon}_{Yi}' \\ \hat{\varepsilon}_{3i}\hat{\varepsilon}_{Yi}' \end{pmatrix} I_i, \tag{2.10}$$

where $\hat{\varepsilon}_{ri} = X_{ri} - \hat{X}_{ri}$ $(1\le r\le 3)$, $\hat{\varepsilon}_{Yi} = Y_i - \hat{Y}_i$ and, for $\zeta = X_1, X_2, X_3, Y$, $\hat{\zeta}_i$ denotes a nonparametric estimate of $E[\zeta_i|Z_i]$. The exact expression of the trimming function I_i and the nonparametric estimate $\hat{\zeta}_i$ will be given below according to the underlying distribution of Z. Now, using the estimate $\hat{\beta}_2$ obtained from (2.10), we can estimate $(\gamma_1,\delta,\beta_3)$ in the partially linear model (2.9) by

$$\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\delta} \\ \hat{\gamma}_3 \end{pmatrix} = \hat{\Gamma}^{-1} n^{-1} \sum_i \begin{pmatrix} \hat{\varepsilon}_{1i}\hat{\varepsilon}_{Yi} \\ \hat{\beta}_2'\hat{\varepsilon}_{2i}\hat{\varepsilon}_{Yi} \\ \hat{\varepsilon}_{3i}\hat{\varepsilon}_{Yi} \end{pmatrix} I_i, \tag{2.11}$$

where

$$\hat{\Gamma} = n^{-1}\sum_i \begin{pmatrix} \hat{\varepsilon}_{1i}\hat{\varepsilon}_{1i}' & \hat{\varepsilon}_{1i}\hat{\varepsilon}_{2i}'\hat{\beta}_2 & \hat{\varepsilon}_{1i}\hat{\varepsilon}_{3i}' \\ \hat{\beta}_2'\hat{\varepsilon}_{2i}\hat{\varepsilon}_{1i}' & \hat{\beta}_2'\hat{\varepsilon}_{2i}\hat{\varepsilon}_{2i}'\hat{\beta}_2 & \hat{\beta}_2'\hat{\varepsilon}_{2i}\hat{\varepsilon}_{3i}' \\ \hat{\varepsilon}_{3i}\hat{\varepsilon}_{1i}' & \hat{\varepsilon}_{3i}\hat{\varepsilon}_{2i}'\hat{\beta}_2 & \hat{\varepsilon}_{3i}\hat{\varepsilon}_{3i}' \end{pmatrix} I_i. \tag{2.12}$$

Finally, the t-ratio which we will use as statistic for our test is

$$t = n^{1/2}\hat{\delta}/(\hat{\sigma}^2\hat{a}_{(k+1)(k+1)})^{1/2}, \tag{2.13}$$

where $\hat{a}_{(k+1)(k+1)}$ is the $(k+1)$th diagonal element in $\hat{\Gamma}^{-1}$, $\hat{\sigma}^2 = \sum_i n^{-1}\hat{U}_i^2 I_i$ and $\hat{U}_i = \hat{\varepsilon}_{Yi} - \hat{\varepsilon}_{1i}'\hat{\gamma}_1 - \hat{\varepsilon}_{2i}'\hat{\beta}_2\hat{\delta} - \hat{\varepsilon}_{3i}'\hat{\gamma}_3$.

2.2. Asymptotic properties

2.2.1. Discrete regressors.

First we suppose that Z is an $\mathbb{R}^q$-valued discrete random variable, i.e.,

$$\exists\ \mathcal{D}\subset\mathbb{R}^q,\ \mathcal{D} \text{ countable set, such that } P(Z\subset\mathcal{D})=1 \text{ and } z_i\in\mathcal{D} \Rightarrow P(Z=z_i)>0. \tag{2.14}$$

This assumption was already discussed in Delgado and Mora (1994). The simplest nonparametric weights we can use in this case are the *non-smoothing weights*, that is to say, for $j\neq i$

$$W_{nj}(Z_i) = I(Z_j=Z_i)/(\textstyle\sum_{c\neq i} I(Z_c=Z_i)), \tag{2.15}$$

where $I(A)$ is the indicator function of event A. We also define

$$I_i = I(\textstyle\sum_{c\neq i} I(Z_c=Z_i)>0). \tag{2.16}$$

This trimming function is introduced in order to consider only those observations for which the denominator in (2.15) is not 0.

In some situations the non-smoothing weights may perform poorly, so that we also consider in this chapter two well-known smoothing weights, namely, the *kernel weights* and the k-*nearest neighbour weights*. The former ones are

$$\tilde{W}_{nj}(Z_i) = \psi((Z_i-Z_j)/h_n)/\sum_{c\neq i}\psi((Z_i-Z_c)/h_n), \tag{2.17}$$

for a kernel function ψ from $\mathbb{R}^q$ to $\mathbb{R}$ and a sequence of smoothing values h_n satisfying that

$$\psi \text{ has bounded support and } h_n \longrightarrow 0 \text{ (as } n \longrightarrow \infty). \tag{2.18}$$

The precise definition of k-nearest neighbour weights is somewhat involved -we refer the interested reader to Stone (1977) for the general case or Delgado and Mora (1994) for the discrete case. We assume that

$$1/k_n + k_n/n \longrightarrow 0. \tag{2.19}$$

In order to establish the asymptotic properties of t we will assume

$$Var(Y|X,Z)=\sigma^2\in(0,\infty),\ E((X-E[X|Z)(X-E[X|Z)')\equiv \Phi \text{ is d.p.} \tag{2.20}$$

Hereafter we will denote $\Phi_{ij} \equiv E((X_i-E[X_i|Z)(X_j-E[X_j|Z)')$. Thus,

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} \\ \Phi'_{12} & \Phi_{22} & \Phi_{23} \\ \Phi'_{13} & \Phi'_{23} & \Phi_{33} \end{bmatrix}. \tag{2.21}$$

The homoskedasticity assumption in (2.20) will be relaxed below. The assumption on Φ is an identifiability condition for the unknown parameter β. The following theorem establishes the asymptotic properties of the test-statistic t under H_0 and H_1.

THEOREM 1.- Let $\{(Y_i,X_{1i},X_{2i},X_{3i},Z_i),\ 1\le i\le n\}$ *be independent identically distributed observations from an* $\mathbb{R}\times\mathbb{R}^k\times\mathbb{R}^l\times\mathbb{R}^m\times\mathbb{R}^q$*-valued observable random variable* (Y,X_1,X_2,X_3,Z) $(k+l+m = p)$. *Assume* that (2.14) and (2.20) hold, $E\|X\|^4<\infty$ and $E[U^4]<\infty$ (where $U\equiv Y-E[Y|X,Z]$), and suppose we use the weights defined in (2.15) and the trimming function defined in (2.16).

a) Under H_0 (i.e. if (2.1) holds), if we denote

$$\Sigma_{23}\equiv\begin{pmatrix}\Phi_{22} & \Phi_{23}\\ \Phi'_{23} & \Phi_{33}\end{pmatrix},\quad \begin{pmatrix}\alpha_2\\ \alpha_3\end{pmatrix}\equiv\Sigma_{23}^{-1}\begin{pmatrix}\Phi'_{12} & \Phi_{23}\\ \Phi'_{13} & \Phi_{33}\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_3\end{pmatrix},$$

and $\varepsilon_r\equiv X_r-E[X_r|Z]$ $(1\le r\le 3)$, then

a1) $$\begin{pmatrix}\hat\beta_2\\ \hat\beta_3\end{pmatrix}\xrightarrow{p}\begin{pmatrix}\alpha_2\\ \alpha_3\end{pmatrix}.$$

a2) If $\alpha_2\neq 0$ then,

$$n^{1/2}\begin{pmatrix}\hat\gamma_1-\beta_1\\ \hat\delta\\ \hat\gamma_3-\beta_3\end{pmatrix}\xrightarrow{d}N(\begin{pmatrix}0\\0\\0\end{pmatrix},\sigma^2\Gamma^{-1}),$$

where $\Gamma\equiv H(\alpha_2)'\Phi H(\alpha_2)$ and $\forall u\in\mathbb{R}^l$,

$$H(u)\equiv\begin{pmatrix}I_k & 0 & 0\\ 0 & u & 0\\ 0 & 0 & I_m\end{pmatrix}.$$

a3) If $\alpha_2\neq 0$ then, $t\xrightarrow{d}N(0,1)$.

b) Under H_1 (i.e. if (2.2) holds), then

b1) $$n^{1/2}\begin{pmatrix}\hat\beta_2-\beta_2\\ \hat\beta_3-\beta_3\end{pmatrix}\xrightarrow{d}N(\begin{pmatrix}0\\0\end{pmatrix},\sigma^2\Sigma_{23}^{-1}).$$

b2) If $\beta_2\neq 0$ then,

$$\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\delta} \\ \hat{\gamma}_3 \end{pmatrix} \xrightarrow{p} \begin{pmatrix} 0 \\ 1 \\ \beta_3 \end{pmatrix}.$$

b3) If $\beta_2 \neq 0$ then, $\forall\ \rho > 0$ $P(|t| > \rho) \longrightarrow 1$ (as $n \longrightarrow \infty$). ■

COROLLARY 1.- All results stated in theorem 1 also hold when

a) kernel weights (2.17) *and trimming function* (2.16) *are used and we also assume* (2.18).

b) k-NN weights are used and we also assume (2.19). *(No trimming function is required here).* ■

The homoskedasticity assumption in (2.20) may be suppressed, but then all asymptotic variance-covariance matrices change in the usual way (see Eicker 1963, White 1980). Specifically, if instead of (2.20) we assume that

$$Var(Y|X,Z) = \sigma^2(X,Z) > 0 \text{ and } E((X-E[X|Z)(X-E[X|Z)') \equiv \Phi \text{ is d.p.} \tag{2.22}$$

then, according to (2.6), the statistic which we must use is

$$t^* = n^{1/2}\hat{\delta}/(\hat{b}_{(k+1)(k+1)})^{1/2}, \tag{2.23}$$

where $\hat{b}_{(k+1)(k+1)}$ is the $(k+1)$th diagonal element in $\hat{\Gamma}^{-1}\hat{\Psi}\hat{\Gamma}^{-1}$ and

$$\hat{\Psi} = n^{-1}\sum_i \begin{pmatrix} \hat{\varepsilon}_{1i}\hat{\varepsilon}'_{1i} & \hat{\varepsilon}_{1i}\hat{\varepsilon}'_{2i}\hat{\beta}_2 & \hat{\varepsilon}_{1i}\hat{\varepsilon}'_{3i} \\ \hat{\beta}'_2\hat{\varepsilon}_{2i}\hat{\varepsilon}'_{1i} & \hat{\beta}'_2\hat{\varepsilon}_{2i}\hat{\varepsilon}'_{2i}\hat{\beta}_2 & \hat{\beta}'_2\hat{\varepsilon}_{2i}\hat{\varepsilon}'_{3i} \\ \hat{\varepsilon}_{3i}\hat{\varepsilon}'_{1i} & \hat{\varepsilon}_{3i}\hat{\varepsilon}'_{2i}\hat{\beta}_2 & \hat{\varepsilon}_{3i}\hat{\varepsilon}'_{3i} \end{pmatrix} \hat{U}_i^2 I_i.$$

The following theorem summarises the asymptotic properties of this heteroskedasticity consistent t-ratio t^*. Corresponding results for $(\hat{\beta}_2, \hat{\beta}_3)$ and $(\hat{\gamma}_1, \hat{\delta}, \hat{\gamma}_3)$ under H_0 and H_1 may be obtained in a similar way for this heteroskedastic model.

COROLLARY 2.- With the same conditions and notation as in Theorem 1, let us replace assumption (2.20) *by assumption* (2.22).

a) *If* H_0 *is true and* $\alpha_2 \neq 0$, *then* $t^* \xrightarrow{d} N(0,1)$.

b) *If* H_1 *is true and* $\beta_2 \neq 0$, *then* $\forall\ \rho > 0$, $P(|t| > \rho) \longrightarrow 1$. ■

2.2.2. Mixed continuous-discrete regressors.

Now we consider the case when Z also contains absolutely continuous random variables, i.e., we will assume that

$$\left.\begin{array}{l} Z = (Z^{(1)}, Z^{(2)}),\ \text{where } Z^{(1)} \subset \mathbb{R}^s \text{ is discrete and} \\ Z^{(2)} \subset \mathbb{R}^r \text{ is absolutely continuous; } r+s = q,\ r \geq 1. \end{array}\right\} \qquad (2.24)$$

We estimate $m_{\zeta l} \equiv E[\zeta_l | Z_l]$ using Nadaraya-Watson kernel weights (Nadaraya 1964, Watson 1964) for the continuous regressors and the non-smoothing weights for the discrete regressors, i.e.

$$W_{nj}(Z_l) = K_{lj}(h_n) I(Z_l^{(1)} = Z_j^{(1)}) / \textstyle\sum_c K_{lc}(h_n) I(Z_l^{(1)} = Z_c^{(1)}), \qquad (2.25)$$

where hereafter we denote $K_{lj}(h_n) \equiv K((Z_l^{(2)} - Z_j^{(2)})/h_n)$, K is a function from $\mathbb{R}^r$ to $\mathbb{R}$ defined by $K(z) = k(z_1)k(z_2)\cdots k(z_q)$, k is a function from $\mathbb{R}$ to $\mathbb{R}$ ("kernel function") and h_n is a sequence of positive real numbers which we will refer to as "sequence of smoothing values". Now, we can estimate $m_{\zeta l}$ by $\hat{m}_{\zeta l} = \sum_j \zeta_j W_{nj}(Z_l)$, for any random variable ζ. Note that, unlike in previous sections, this is not a "leave-one-out" estimator. Using these estimates it is possible to construct estimated residuals $\hat{\varepsilon}_{\zeta l}$ and estimates of the parameters of interest $\hat{\Phi}$, $\hat{\beta}$ and $\hat{\sigma}^2$ as in the discrete case, but now

$$I_l = I(\textstyle\sum_k K_{lk}(h_n) I(Z_l^{(1)} = Z_k^{(1)})/nh_n^q > b_n), \qquad (2.26)$$

where b_n is a sequence of positive real numbers (trimming values).

Some additional assumptions are required to prove that a similar result to Theorem 1 holds when there are both continuous and discrete regressors in the unknown part of the model. Given $\gamma \in \mathcal{D}$, let $f_\gamma(.)$ be the density function of $Z^{(2)} | Z^{(1)} = \gamma$ and denote $\theta_\gamma(u) \equiv g(\gamma, u)$ (function $g(.,.)$ as defined in (1.1) for $(\gamma, u) \in \mathbb{R}^q$), $\xi_\gamma(u) \equiv E[X | Z^{(1)} = \gamma, Z^{(2)} = u]$. We will assume that

$$U \equiv Y - E[Y|Z] \text{ and } (X,Z) \text{ are independent,} \tag{2.27}$$

$$\left.\begin{array}{l} \exists v \in \mathbb{N} : f_{\mathfrak{z}} \in \mathcal{G}^{\infty}_{vs},\ \theta_{\mathfrak{z}} \in \mathcal{G}^{4}_{vs}(Z^{(2)}|Z^{(1)}=\mathfrak{z}) \text{ and} \\ \xi_{\mathfrak{z}} \in \mathcal{G}^{2}_{vs}(Z^{(2)}|Z^{(1)}=\mathfrak{z}), \text{ uniformly in } \mathcal{D} \end{array}\right\}, \tag{2.28}$$

$$\text{the kernel function } k \text{ is in the class } \mathcal{K}_{2vs} \text{ and} \tag{2.29}$$

$$b_n \longrightarrow 0, \qquad nb_n^{-4}h_n^{4vs} \longrightarrow 0, \quad nb_n^{4}h_n^{2s} \longrightarrow \infty \ (\text{as } n \longrightarrow \infty) \tag{2.30}$$

Classes $\mathcal{G}^{\alpha}_{\mu}$, $\mathcal{G}^{\infty}_{\mu}$ and $\mathcal{K}_{l}$ are as defined in Robinson (1988), and "uniformly in $\mathcal{D}$" means that the constants which appear in the definition do not depend on the value $\mathfrak{z}$. The following result generalises Theorem 1.

THEOREM 2.- All results stated in Theorem 1 also hold when assumption (2.14) *is replaced by assumption* (2.24), *nonparametric weights* (2.15) *and trimming function* (2.16) *are replaced, respectively, by* (2.25) *and* (2.26), *and, additionally, conditions* (2.27), (2.28), (2.29) *and* (2.30) *are assumed in the model.* ■

Observe that the conditions required in Theorem 2 are much stronger than those required in Theorem 1 -this is why we have preferred to consider first the discrete case separately, as in many cases (see, for example, Section 3 below) all variables in the unknown part of the model are discrete.

The Central Limit Theorems stated in Theorems 1 and 2 generalise earlier results obtained in a complete parametric environment by Davidson and MacKinnon (1981), MacKinnon et al. (1983) and Fisher (1983), among others. Assumptions in Theorem 2 were first introduced by Robinson (1988) in the semiparametric i.i.d. partially linear regression model with absolutely continuous Z, and have been also considered in Delgado and Mora (1994) for the mixed case.

Observe also that the asymptotic distribution of the test-statistic t holds only if the unknown parameters α_2 (under H_0) or β_2 (under H_1) are not 0. Obviously, the assumption $\beta_2 \neq 0$ is not restrictive at all, because if β_2 were equal to 0 (and this hypothesis can be tested), then the alternative hypothesis would be nested in the null hypothesis and the test could be carried out in a straightforward way. The assumption $\alpha_2 \neq 0$ is not restrictive either[3], because, in practice, before facing H_0 and H_1, the researcher estimates separately each model and will only take into account both models if she can accept that the coefficients in them are not 0. (In fact, in view of Theorem 1, in order to make sure that when H_0 is true the percentage of rejections is, asymptotically, not greater than the level of significance, first we should test whether the coefficient of X_2 is 0 or not). We will examine the behaviour of the test-statistic when $\alpha_2 = 0$ in one of the models which we simulate in Section 2.3 below.

2.3. Simulations

We have generated observations from 12 different models in order to examine the finite-sample properties (size and power) of the test. Results corresponding to two different sample sizes are contained in Tables 1, 2, 3 and 4.

In models 1-6 the null hypothesis was true, whereas in models 7-12 the alternative hypothesis was the true one. Except for models 4 and 10, the test we performed was

$$H_0: \ E[Y|X,Z] = X_1'\beta_1 + g(Z) \quad a.s.$$

$$H_1: \ E[Y|X,Z] = X_2'\beta_2 + g(Z) \quad a.s.$$

In models 4 and 10, the null and alternative hypotheses were

$$H_0: \ E[Y|X,Z] = X_1\beta_1 + X_3\beta_3 + g(Z) \quad a.s.$$

$$H_1: \ E[Y|X,Z] = X_2\beta_2 + X_3\beta_3 + g(Z) \quad a.s.$$

---

[3] This assumption is similar to Davidson and MacKinnon's (1981) assumption A3 on the limiting behaviour of $n^{-1}F'(\beta)G(\gamma)$, and appears in all the econometric literature on non-nested tests.

The description of all variables in each model is as follows:

Model 1: $Z\sim Poisson(2)$; V_1,V_2,U all $N(0,1)$; Z,V_1,V_2,U indep.;

$X_1 = 2Z+V_1$; $X_2 = 3Z+V_1+V_2$; $Y = 4X_1+3Z+U$.

Model 2: Identical to model 1, but $Z\sim N(2,2)$.

Model 3: Identical to model 2, but $U = (U^*-2)/2$, with $U^*\sim$CHI2(2).

Model 4: $Z\sim Poisson(2)$; V_1,V_2,V_3,U all $N(0,1)$;

Z,V_1,V_2,V_3,U indep.; $X_1 = 2Z+V_1$; $X_2 = 3Z+V_1+V_2$;

$X_3 = Z+V_1+V_2+V_3$; $Y = 4X_1+X_3+3Z+U$.

Model 5: $Z\sim Poisson(2)$; V_1,V_2,V_3,U all $N(0,1)$;

Z,V_1,V_2,V_3,U indep.; $X_1' = (2Z+V_1,Z+V_1+V_2+V_3)$;

$X_2 = 3Z+V_1+V_2$; $Y = (4,-2)X_1+3Z+U$.

Model 6: $Z_1\sim N(50.5,228.2)$; $Z_2 = Z_2^*+1$, with $Z_2^*\sim Poisson(2.5)$;

$X_1' = (V,V^2)$; $X_2' = (exp(V),exp(-V))$, with $V\sim N(13.4,0.52)$;

$Y = (0.229,-0.0148)X_1+g(Z_1,Z_2)+U$, with $U\sim N(0,0.01^2)$;

Z_1,Z_2^*,V,U indep. $g(Z_1,Z_2) = \left\{ \begin{array}{ll} -0.1 & if\ Z_2 < 2 \\ 0.1 & if\ Z_2 > 3 \\ 0 & otherwise \end{array} \right\}$.

Model 7: Identical to Model 1, but $Y = 0.4X_2 + 3Z + U$.

Model 8: Identical to Model 2, but $Y = 0.4X_2 + 3Z + U$.

Model 9: Identical to Model 3, but $Y = 0.6X_2 + 3Z + U$.

Model 10: Identical to Model 4, but $Y = 0.8X_2 + X_3 + 3Z + U$.

Model 11: Identical to Model 5, but $Y = X_2 + 3Z + U$.

Model 12: Identical to Model 6, but $Y = (c_1,c_2)X_2 + g(Z_1,Z_2) + U$,

where $c_1= -1.3E-7$, $c_2= 1.6E+4$.

In models 1, 2, and 3 the competing regression functions have not any common independent variable (X_3), and we consider the cases of discrete Z (model 1), continuous Z (models 2 and 3) and non-normal error term (model 3).

In model 4 there is an independent variable (X_3) which appears in both the null hypothesis and the alternative one. Model 5 has been defined in such a way that $\alpha_2 = 0$ -thus, the asymptotic distribution for the t-ratio t proved in Theorem 1 does not hold. Finally, model 6 has been defined in such a seemingly strange way in order to make the generated data similar to the observations which we use in Section 3. In this model, the mean and variance of V coincide with the sample mean and variance of $log(X)$, where X is the variable total expenditure considered in Section 3; vector β_1 coincides with the semiparametric estimate of this parameter obtained in Section 3 (see Equation 3 in Table 6); and the variance of U was selected in such a way that the variance of variable Y in this model coincides with the variance of variable Y in Section 3 (percentage of total expenditure spent on food). Observe that, in model 6, functions $\theta_{\gamma}(.)$ and $\xi_{\gamma}(.)$ are both constants, so that this model satisfies (2.28).

Models 7-12 have been generated in a similar way to models 1-6 (respectively), but now the true specification for the regression function is the alternative hypothesis. In model 12, vector β_2 coincides with the semiparametric estimate of this parameter obtained in Section 3 (see Equation 9 in Table 7).

We have used two different sample sizes ($n= 40$ in tables 1 and 3, $n=200$ in tables 2 and 4) with different number of replications in each case ($r=10000$ for $n=40$ and $r=2000$ for $n=200$). In those models in which Z was a Poisson random variable we computed the test-statistic t using a non-smoothing estimate (row labelled Non-S.) and a kernel estimate. In those models in which Z was an absolutely continuous random variable t was computed using a nonparametric kernel estimate with two different smoothing values. In models 6 and 12 we used a kernel estimate for Z_1 and non-smoothing weights for Z_2. All programs were written in FORTRAN double precision and all observations were generated using the pseudorandom number generators from the NAGLIB library. Programs were run at the VAX work-stations in *London School of Economics*.

In all kernel estimates the kernel function we used is the univariate Epanechnikov kernel or the product of univariate Epanechnikov kernels. On computing the kernel estimates smoothing values (h_n) had to be selected; in all cases, we have simply selected meaningful values for h_n –in fact, results were not very sensitive to the choice of this number, provided that it was chosen within a sensible interval. In models with continuous Z we also had to select trimming values: we used $b= 0.1$ for $n= 40$ and $b= 0.03$ for $n= 200$.

In all tables we report the empirical significance level for two nominal significance levels (0.05 and 0.01). We also report the sample mean (M) and variance (V) of the estimate $(\hat{\gamma}_1,\hat{\delta})$ computed as defined in (2.11) (in models 4 and 10 the estimate is $(\hat{\gamma}_1,\delta,\hat{\gamma}_3)$). In small numbers we report the parameters of the normal distribution which asymptotically approximates these estimates according to Theorem 1 (for example, in model 1, $n^{1/2}(\hat{\gamma}_1-4) \xrightarrow{d} N(0,2)$; thus $\hat{\gamma}_1$ asymptotically behaves as a random variable with distribution $N(4,2/n)$).

In tables 1 and 2 we analyse the size of the test. We observe that for a small size (Table 1) the empirical level is always greater than the nominal significance level; however, with medium sample size (table 2) the empirical level and the nominal significance level are very close. The only exceptions to this are the kernel estimates in models 1 and 4, in which, due to the smoothing, the estimate $\hat{\delta}$ is highly biased and, hence, the test does not work (the problem of smoothing in nonparametric and semiparametric regression with discrete variables has already been analysed in Delgado and Mora 1994). Finally, the results obtained in model 5 suggest that when $\alpha_2=0$ the test-statistic may still work, even though the asymptotic result for $\hat{\delta}$ does not hold any more (note that in model 5 the variance of $\hat{\delta}$ does not decrease as the sample size increases).

In tables 3 and 4 we analyse the power of the test. The proposed test is extremely powerful even for a small sample size. Only in models 1 and 2 the percentage of rejections for $\alpha=0.05$ is below 75%.

To sum up, in this small Monte Carlo experiment the proposed test performs adequately with respect to its size and power characteristics, though with small sample size there is a tendency to over-reject. Thus, the test-statistic behaves in a similar way to other test-statistics based on AN procedures in non-nested parametric and semiparametric models (see Pesaran 1982, Godfrey and Pesaran 1983 or Delgado and Stengos 1994, among others).

3. TESTING FUNCTIONAL FORM OF ENGEL CURVES

3.1. The Engel Curve relationship

The traditional static Engel Curve model specifies the relationship between total expenditure of a household and expenditure on a specific good s,

$$p_s q_s = f_s(X,Z) + U, \tag{3.1}$$

where p_i is price of good i (which is assumed to be constant for all households), q_i is quantity purchased of good i, X is total expenditure of the household and Z is a vector of other possibly relevant variables.

An alternative model which has been widely studied is the Engel Curve relationship in share form, i.e.,

$$Y_s = g_s(X,Z) + U, \tag{3.2}$$

where $Y_i = p_i q_i / X$ denotes the proportion of total expenditure spent on good i. The interest on this share form of the Engel Curve relationship started with the studies of Working (1943) and Leser (1963), who revealed the stability of the log-linear specification for Food Engel Curves,

$$Y_s = \beta_0 + \beta_1 log(X) + U. \tag{3.3}$$

However, this simple formulation does not provide a reasonable approximation for other Engel Curves (see, for example, Atkinson et al. 1990 or Lewbel 1991, and references therein) and the class of preferences that can generate (3.3) (which has been called PIGLOG) underlies the Almost Ideal model of Deaton of Muellbauer, which is not consistent with empirical evidence on the shape of Engel Curves (see Lewbel 1991 or Banks et al. 1993). In recent studies, the Working-Leser form of Engel Curves has been generalised in different ways in order to obtain global demand systems consistent with empirical evidence. Two of the most popular specifications for static Engel Curves are the generalised form of the Working-Leser relationship

$$Y_s = \beta_0 + \beta_1 log(X) + \beta_2 log(X)^2 + U, \tag{3.4}$$

and the share form relationship for Engel Curves derived from the Quadratic Expenditure System (QES) of Pollack and Wales (1978, 1980),

$$Y_s = \beta_0 + \beta_1 X + \beta_2 X^{-1} + U. \tag{3.5}$$

In this section we want to face the functional forms for Engel Curves (3.4) and (3.5) taking into account other possibly relevant variables. Specifically, we consider a vector of variables Z containing what has usually been referred to as "*demographic variables*" (the four demographic variables considered in this paper are described in Section 3.2 below). The traditional procedures commonly used to incorporate the information of these variables in the modelling of Engel Curves are demographic translating and demographic scaling (see Pollack and Wales 1981). These procedures are simple to implement but recent empirical research does not support this specification (see Gozalo 1992). In this paper we use a different method to model the effect of demographic variables: we have generalised equations (3.4) and (3.5) including a vector Z of other possible relevant variables in a semiparametric way, yielding the stochastic specification

$$E[Y_s|X,Z] = \beta_0 + \beta_1 log(X) + \beta_2 log(X)^2 + g(Z) \quad a.s., \tag{3.6}$$

$$E[Y_s|X,Z] = \beta_0 + \beta_1 X + \beta_2 X^{-1} + g(Z) \quad a.s., \tag{3.7}$$

for the generalised Working-Leser form (hereafter referred to as "Working-Leser form") and the Quadratic Expenditure form (hereafter referred to as "QES" form) respectively.

3.2. Data

Our results are based on data from the 1980 Consumer Expenditure Survey (CES) for Spain. The CES contains *23972* observations, with detailed information on household characteristics, total income and expenditure and various expenditure categories. The sample is designed to be representative of the total Spanish population.

The main variables we have considered in our analysis are:

- X: Total expenditure (in pesetas),
- Y: Percentage of total expenditure spent on food ("share food"),
- Z_1: Age of "reference person" in the household (i.e. member of the household with greatest income),
- Z_2: Size of household,
- Z_3: Size of the town where the household is placed,
- Z_4: Sex of reference person.

The variable Z_3 was categorised into *5* groups according to the number of inhabitants (in thousands) of the town (NI): $Z_3(NI) = i$ if $NI \in I_i$, where $I_1=(0,2)$, $I_2=[2,10)$, $I_3=[10,50)$, $I_4=[50,200)$, $I_5=[200,\infty)$.

In this illustrative study we have restricted our analysis to food Engel Curves in order to avoid problems with zeros (in all other expenditure groups there was a meaningful proportion of observations exactly equal to *0*).

Table 5 provides a descriptive summary of the most important variables. Figure 1 shows the nonparametric density estimate of X. This plot shows that the lognormal specification may be appropriate for this variable (in all kernel estimates we used univariate Epanechnikov kernels or the product of univariate Epanechnikov kernels). This nonparametric estimate was not computed for values of X greater than *2.3E+6* because the number of observations decreased in that interval.

Figure 2 is a cross-plot of Y and X with all observations in the sample. Figures 3a, 3b, 4 and 5 depict the estimated Engel Curves in share form. In figures 3a and 3b we used the whole data set to obtain the nonparametric estimates. These figures show that the log-linear specification may be appropriate for this regression function. In figures 4 and 5 we tried to examine the influence of other variables in the shape of food Engel Curves. First we analysed the influence of the variable "age of reference person": we estimated separately Engel Curves for three different types of households (type A if $Z_1 \le 30$, type B if $Z_1 \in (30,60]$ and type C if $Z_1 > 60$) -the corresponding estimates are shown in figure 3. We observe that the shape of the three estimated Engel Curves is quite similar. Finally, we estimated separately Engel Curves for households of size *1*, *2-5* and *6 or +*. Apparently, this variable (Z_2) has more influence on the shape of Engel Curves than Z_1: in figure 4 we note that the form of the estimated Engel Curve for households with size *1* seems to differ from the form which has been estimated for households with size greater than *1*. In figures 3, 4 and 5 we observe that the Engel Curve behaves in a strange way when X is large (see, for instance, the shapes of Food Engel Curves for households of type C in figure 3 or for households with size 1 in Figure 4). This is possibly due to the fact that the observations of X are more sparse when X is large and, hence, the nonparametric estimate has high variance in that interval (it would be necessary to use a bigger smoothing value in order to reduce the variance of the nonparametric estimate when X is large).

3.3. Estimation procedure and results

The objective of this section is to use the test which has been proposed in Section 2 to face two competing explanatory models for the Engel Curve relationship: the Working-Leser form (3.6) and the QES form (3.7) described in Section 3.1 above.

First of all we have estimated semiparametrically different regression functions trying to analyse the importance of different characteristics of the households on estimating the parameters of the Engel Curve form.

The first six equations we have estimated are derived from the Working-Leser form. If we denote $W_1' = (logX, (logX)^2)$ then we have

<u>Eq. 1</u>: $E[Y|W_1,Z] = \beta' W_1 + \beta_0$

<u>Eq. 2</u>: $E[Y|W_1,Z] = \beta' W_1 + g(Z_1)$

<u>Eq. 3</u>: $E[Y|W_1,Z] = \beta' W_1 + g(Z_2)$

<u>Eq. 4</u>: $E[Y|W_1,Z] = \beta' W_1 + g(Z_1,Z_2)$

Eq. 5: $E[Y|W_1,Z] = \beta'W_1 + g(Z_1,Z_2,Z_3)$

Eq. 6: $E[Y|W_1,Z] = \beta'W_1 + g(Z_1,Z_2,Z_3,Z_4)$

Equations 7-12 are derived from the Q.E.S. form. Specifically, if we denote $W_2'=(X, X^{-1})$, then we have

Eq. 7: $E[Y|W_2,Z] = \beta'W_2 + \beta_0$

Eq. 8: $E[Y|W_2,Z] = \beta'W_2 + g(Z_1)$

Eq. 9: $E[Y|W_2,Z] = \beta'W_2 + g(Z_2)$

Eq.10: $E[Y|W_2,Z] = \beta'W_2 + g(Z_1,Z_2)$

Eq.11: $E[Y|W_2,Z] = \beta'W_2 + g(Z_1,Z_2,Z_3)$

Eq.12: $E[Y|W_2,Z] = \beta'W_2 + g(Z_1,Z_2,Z_3,Z_4)$

Equations 1 and 7 were estimated using the OLS procedure, whereas in all other equations $\hat{\beta}$ was computed using the semiparametric estimate defined in (2.4). We used non-smoothing weights for Z_2, Z_3, Z_4 and kernel weights with two different smoothing values for Z_1 (these values may seem too big at first sight, but remember that the sample mean and standard deviation of Z_1 are, respectively, *50* and *15*). In table 6 we report the semiparametric estimates which have been obtained for eqs. 1-6 and in table 7 we report results for eqs. 7-12. In all equations both parameters are significantly different from *0* irrespective of whether we assume homoskedasticity or not ($\alpha=0.05$). As regards the importance of other variables, Z_2 seems to be the one with greater influence on results: note that the estimates of models 1-2 (models 7-8) are similar, but they seem to differ from those of model 3 (model 9).

We have performed 12 different tests in order to face our competing hypothesis taking into account the influence of those variables which describe households. In tests 1-6 the null hypothesis is the Working Leser form and the alternative is the Q.E.S. form, i.e.,

Tests 1-6:
$$\left.\begin{array}{l} H_0\text{: } E[Y|W_1,W_2,Z] = \beta'W_1 + g(Z) \\ H_1\text{: } E[Y|W_1,W_2,Z] = \beta'W_2 + g(Z) \end{array}\right\}.$$

In tests 7-12 we reversed the null and alternative hypotheses, i.e.,

$$\text{Tests 7-12:} \qquad \left.\begin{aligned} H_0&: E[Y|W_1,W_2,Z] = \beta' W_2 + g(Z) \\ H_1&: E[Y|W_1,W_2,Z] = \beta' W_1 + g(Z) \end{aligned}\right\} .$$

The set of variables included in Z was different for each test:

Tests 1 and 7:	No Z variable;
Tests 2 and 8:	$Z = Z_1$;
Tests 3 and 9:	$Z = Z_2$;
Tests 4 and 10:	$Z = (Z_1, Z_2)$;
Tests 5 and 11:	$Z = (Z_1, Z_2, Z_3)$;
Tests 6 and 12:	$Z = (Z_1, Z_2, Z_3, Z_4)$.

Note that with these definitions, tests 1 and 7 are purely parametric but all other tests are semiparametric.

In table 8 we report the results obtained for all tests, using the whole data set. In all cases we reject the null hypothesis (an asterisk at the end of the row means that the null hypothesis is rejected with significance level $\alpha=0.05$). However, results for tests 1-6 are substantially different from those obtained for tests 7-12. From our estimation of tests 1-6, we deduce that the value of δ is significantly different from 0, but also significantly different from 1 -thus, we must reject the null hypothesis, but it is also unreasonable to think that the alternative one is correct. On estimating the parameters involved in tests 7-12, we also obtain that δ is significantly different from 0 and 1, but, at least, $\hat{\delta}$ seems to be close to 1 (if H_1 were true, $\hat{\delta}$ would converge to 1 in probability). In short, we cannot accept either the Working-Leser form or the Q.E.S. form as the true underlying forms for the Engel Curves. As mentioned in Section 2, the test we propose tends to over-reject in some situations. In order to analyse whether this could be the case in our model, we generated models 6 and 12 in Section 2.3. Those models were generated in such a way that the resulting variables, individually, had similar distribution to the variables considered in this model. The problem of over-rejection did not appear in those simulated models for sample sizes much smaller than the one we have in this real data problem. Thus, it also seems unreasonable to say that the results obtained in table 8 are a consequence of the bad performance of the test-statistic we used.

We have searched for explanations for these negative results. First of all we have reduced the data set. It might happen that the rejectionof both the Working-Leser and the QES form is a consequence of considering too heterogeneous a sample. In order to analyse this conjecture, we performed again all tests but now considering only those observations corresponding to "standard" households, i.e., those households consisting of two adults (one woman and one man) between 18 and 64 years of age and one, two or three children below 18. The sample size decreased then from *23972* to *6710*, but results (shown in Table 9) were entirely similar to those we had previously obtained. The only remarkable difference was that now the variable "size of household" was not as relevant in these tests as in those performed with the whole data set, that is to say, results are similar for tests 3 and 1-2 in Table 9, in opposition to what is observed in Table 8 (this difference was entirely foreseeable because in the new data set the size of all households was 3, 4 or 5).

Finally, we reduced the data set in a different way: we performed again all tests considering only those households whose total expenditure was within the $(0.1,0.8)$ quantiles. In Table 10 we report the results we obtained. In this table we observe that, for the first time, in some of the tests the null hypothesis was not rejected (tests 1, 2 and 3). Moreover, in tests 1-6, the parameter β is significantly different from 0 (in all cases, significance level $\alpha=0.05$), whereas in tests 7-12 the parameter δ was not significantly different from 1 and, in some cases, the parameter β was not significantly different from 0. Thus, the Working-Leser form adequately explains all results. But this is by no means a surprise. It is well-known that the Working-Leser form adequately explains the relationship between total expenditure (X) and share food (Y), except for those observations contained in the upper tail of X (this is consistent with results discussed, among others, by Banks et al. 1993) -and those are precisely the observations which we did not take into account in these final tests.

To sum up, neither the Working-Leser nor the QES form can be accepted as suitable forms for the Food Engel Curve when all observations in the Spanish Family Expenditure Survey of 1980 are considered. This negative conclusion holds throughout all the semiparametric specifications we considered. The Working-Leser form seems to explain adequately the behaviour of those observations which are not on the tails of the distribution of total expenditure, but fails to explain the results we obtained even when only homogeneous households were considered. This negative result does not seem to arise from the problem of heteroskedasticity or insufficient number of observations (heteroskedastic consistent standard errors were computed and the sample size was quite large).

In fact, if we examine again Figure 2, we might conjecture that the reason why we have obtained this negative result is because there does not seem to be mean dependence between Y and X for those households with high total expenditure (specifically, in Figure 2 we see that when X is greater than *5E+6* the cross-plot between Y and X does not show any relationship between both variables). Our testing procedure discriminates between the two proposed specifications (in favour of the Generalised Working-Leser form) if we do not take into account observations on the tails of the distribution of X. But it fails to discriminate between them when the whole data set is used, possibly because we might admit that there is a structural break in the true regression function: the Generalised Working-Leser form explains the behaviour of households which do not have high total expenditure, but no relationship between Y, X could be proposed for households with high total expenditure[4]. Hence, a possible next step in this research would be to estimate the point at which there is a structural break.

APPENDIX.- Proofs

We will only prove Theorem 1. Corollaries 1, 2 and Theorem 2 follow applying similar arguments as in Theorem 1 and the following results from Delgado and

[4] Obviously, we mean that the only reasonable functional form which might capture the relationship between Y and X (for observations with high total expenditure) would be a constant function not depending on X.

Mora (1994): Corollary 3 (for Corollary 1a), Corollary 4 (for Corollary 1b), Theorem 5 (for Corollary 2) and Theorem 6 (for Theorem 2).

Proof of Theorem 1, a1.- Let us define

$$\begin{pmatrix} \tilde{\beta}_2 \\ \tilde{\beta}_3 \end{pmatrix} = \left\{ \sum_i \begin{pmatrix} \varepsilon_{2i}\varepsilon'_{2i} & \varepsilon_{2i}\varepsilon'_{3i} \\ \varepsilon_{3i}\varepsilon'_{2i} & \varepsilon_{3i}\varepsilon'_{3i} \end{pmatrix} I_i \right\}^{-1} \sum_i \begin{pmatrix} \varepsilon_{2i}\varepsilon_{Yi} \\ \varepsilon_{3i}\varepsilon_{Yi} \end{pmatrix} I_i .$$

It will suffice to prove that

$$\begin{pmatrix} \tilde{\beta}_2-\alpha_2 \\ \tilde{\beta}_3-\alpha_3 \end{pmatrix} \xrightarrow{p} 0, \tag{A.1}$$

and

$$\begin{pmatrix} \hat{\beta}_2-\tilde{\beta}_2 \\ \hat{\beta}_3-\tilde{\beta}_3 \end{pmatrix} \xrightarrow{p} 0. \tag{A.2}$$

We will first prove (A.1).

$$n^{-1}\sum_i \begin{pmatrix} \varepsilon_{2i}\varepsilon'_{2i} & \varepsilon_{2i}\varepsilon'_{3i} \\ \varepsilon_{3i}\varepsilon'_{2i} & \varepsilon_{3i}\varepsilon'_{3i} \end{pmatrix} \xrightarrow{p} \Sigma_{23} \text{ (by LGN);} \tag{A.3}$$

$$n^{-1}\sum_i \begin{pmatrix} \varepsilon_{2i}\varepsilon'_{2i} & \varepsilon_{2i}\varepsilon'_{3i} \\ \varepsilon_{3i}\varepsilon'_{2i} & \varepsilon_{3i}\varepsilon'_{3i} \end{pmatrix} (1-I_i) \xrightarrow{p} 0 \tag{A.4}$$

(applying Cauchy-Schwartz inequality as in Delgado and Mora 1994, because $E\|X\|^4<\infty$). As $\varepsilon_{Yi} = \varepsilon'_{1i}\beta_1 + \varepsilon'_{3i}\beta_3 + U_i$, then

$$n^{-1}\sum_i \begin{pmatrix} \varepsilon_{2i} \\ \varepsilon_{3i} \end{pmatrix} \varepsilon_{Yi} I_i = n^{-1}\sum_i \begin{pmatrix} \varepsilon_{2i}\varepsilon'_{1i} & \varepsilon_{2i}\varepsilon'_{3i} \\ \varepsilon_{3i}\varepsilon'_{1i} & \varepsilon_{3i}\varepsilon'_{3i} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_3 \end{pmatrix} I_i + n^{-1}\sum_i \begin{pmatrix} \varepsilon_{2i} \\ \varepsilon_{3i} \end{pmatrix} U_i I_i .$$

Removing the trimming function as before, we obtain

$$n^{-1}\sum_i \begin{pmatrix} \varepsilon_{2i} \\ \varepsilon_{3i} \end{pmatrix} \varepsilon_{Yi} I_i \xrightarrow{p} \begin{pmatrix} \Phi'_{12} & \Phi_{23} \\ \Phi'_{13} & \Phi_{33} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_3 \end{pmatrix}. \tag{A.5}$$

Thus, (A.1) follows from (A.3), (A.4) and (A.5). As for (A.2),

$$\begin{pmatrix} \hat\beta_2-\tilde\beta_2 \\ \hat\beta_3-\tilde\beta_3 \end{pmatrix} = \hat\Sigma_{23}^{-1}\, n^{-1}\textstyle\sum_1 (\begin{pmatrix} \hat\varepsilon_{21} \\ \hat\varepsilon_{31} \end{pmatrix} \hat\varepsilon_{Y1} I_1 - \Sigma_{23}^{-1}\, n^{-1}\sum_1 (\begin{pmatrix} \varepsilon_{21} \\ \varepsilon_{31} \end{pmatrix} \varepsilon_{Y1} I_1 =$$

$$\hat\Sigma_{23}^{-1}\, n^{-1}\{\textstyle\sum_1 (\begin{pmatrix} \hat\varepsilon_{21} \\ \hat\varepsilon_{31} \end{pmatrix} \hat\varepsilon_{Y1} I_1 - \sum_1 (\begin{pmatrix} \varepsilon_{21} \\ \varepsilon_{31} \end{pmatrix} \varepsilon_{Y1} I_1\} + \{\hat\Sigma_{23}^{-1} - \Sigma_{23}^{-1}\}\, n^{-1}\sum_1 (\begin{pmatrix} \varepsilon_{21} \\ \varepsilon_{31} \end{pmatrix} \varepsilon_{Y1} I_1.$$

Here, the second term converges to 0 in probability because $\{\hat\Sigma_{23}^{-1} - \Sigma_{23}^{-1}\}$ converges to 0 in probability (as proved in Delgado and Mora 1994) and the other factor has already been analised as (A.5). As for the first term, we only have to prove that

$$n^{-1}\{\textstyle\sum_1 (\begin{pmatrix} \hat\varepsilon_{21} \\ \hat\varepsilon_{31} \end{pmatrix} \hat\varepsilon_{Y1} I_1 - \sum_1 (\begin{pmatrix} \varepsilon_{21} \\ \varepsilon_{31} \end{pmatrix} \varepsilon_{Y1} I_1\} \xrightarrow{p} 0.$$

Now, as H_0 is true

$$\varepsilon_{Y1} = \varepsilon'_{11}\beta_1 + \varepsilon'_{31}\beta_3 + U_1,$$

$$\hat\varepsilon_{Y1} I_1 = \hat\varepsilon'_{11}\beta_1 I_1 + \hat\varepsilon'_{31}\beta_3 I_1 + \hat\varepsilon_{U1} I_1.$$

(The second equality follows from equation 3.8 in Delgado and Mora 1994). Therefore,

$$n^{-1}\{\textstyle\sum_1 (\begin{pmatrix} \hat\varepsilon_{21} \\ \hat\varepsilon_{31} \end{pmatrix} \hat\varepsilon_{Y1} I_1 - \sum_1 (\begin{pmatrix} \varepsilon_{21} \\ \varepsilon_{31} \end{pmatrix} \varepsilon_{Y1} I_1\} =$$

$$n^{-1}\{\textstyle\sum_1 \begin{pmatrix} \hat\varepsilon_{21}\hat\varepsilon'_{11} & \hat\varepsilon_{21}\hat\varepsilon'_{31} \\ \hat\varepsilon_{31}\hat\varepsilon'_{11} & \hat\varepsilon_{31}\hat\varepsilon'_{31} \end{pmatrix} I_1\} \begin{pmatrix} \beta_1 \\ \beta_3 \end{pmatrix} + n^{-1}\sum_1 (\begin{pmatrix} \hat\varepsilon_{21} \\ \hat\varepsilon_{31} \end{pmatrix} \hat\varepsilon_{U1} I_1 +$$

$$- n^{-1}\{\textstyle\sum_1 \begin{pmatrix} \varepsilon_{21}\varepsilon'_{11} & \varepsilon_{21}\varepsilon'_{31} \\ \varepsilon_{31}\varepsilon'_{11} & \varepsilon_{31}\varepsilon'_{31} \end{pmatrix} I_1\} \begin{pmatrix} \beta_1 \\ \beta_3 \end{pmatrix} - n^{-1}\sum_1 (\begin{pmatrix} \varepsilon_{21} \\ \varepsilon_{31} \end{pmatrix} U_1 I_1.$$

As before, we can ignore the trimming function. Now, the fourth term converges to 0 by LGN; the second term converges to 0 as is proved in Delgado and Mora 1994; and the difference between the first term and the third one converges in probability to 0 because $\hat\Phi$ converges to Φ (also proved in Delgado and Mora 1994). ■

Proof of Theorem 1, a2.-

$$\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\delta} \\ \hat{\gamma}_3 \end{pmatrix} = \hat{\Gamma}^{-1} n^{-1} \sum_i \begin{pmatrix} \hat{\varepsilon}_{1i} \\ \hat{\beta}_2' \hat{\varepsilon}_{2i} \\ \hat{\varepsilon}_{3i} \end{pmatrix} \hat{\varepsilon}_{Yi} I_i, \tag{A.6}$$

where $\hat{\Gamma} \equiv H(\hat{\beta}_2)'\hat{\Phi}H(\hat{\beta}_2)$, and $\hat{\beta}_2$, $\hat{\Phi}$ are as before. As H_0 is true,

$$\hat{\varepsilon}_{Yi} I_i = \hat{\varepsilon}_{1i}' \beta_1 I_i + \hat{\varepsilon}_{3i}' \beta_3 I_i + \hat{\varepsilon}_{Ui} I_i = (\hat{\varepsilon}_{1i}' \ \ \hat{\varepsilon}_{2i}' \hat{\beta}_2 \ \ \hat{\varepsilon}_{3i}') \begin{pmatrix} \beta_1 \\ 0 \\ \beta_3 \end{pmatrix} I_i + \hat{\varepsilon}_{Ui} I_i .$$

Thus, taking this expression into (A.6) we obtain:

$$\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\delta} \\ \hat{\gamma}_3 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ 0 \\ \beta_3 \end{pmatrix} + \hat{\Gamma}^{-1} n^{-1} \sum_i \begin{pmatrix} \hat{\varepsilon}_{1i} \\ \hat{\beta}_2' \hat{\varepsilon}_{2i} \\ \hat{\varepsilon}_{3i} \end{pmatrix} \hat{\varepsilon}_{Ui} I_i \quad \Rightarrow$$

$$n^{1/2} \begin{pmatrix} \hat{\gamma}_1 - \beta_1 \\ \hat{\delta} \\ \hat{\gamma}_3 - \beta_3 \end{pmatrix} = \hat{\Gamma}^{-1} H(\hat{\beta}_2)' \times n^{-1/2} \sum_i \begin{pmatrix} \hat{\varepsilon}_{1i} \\ \hat{\varepsilon}_{2i} \\ \hat{\varepsilon}_{3i} \end{pmatrix} \hat{\varepsilon}_{Ui} I_i .$$

Now,

$$n^{-1/2} \sum_i \begin{pmatrix} \hat{\varepsilon}_{1i} \\ \hat{\varepsilon}_{2i} \\ \hat{\varepsilon}_{3i} \end{pmatrix} \hat{\varepsilon}_{Ui} I_i \xrightarrow{d} N(0, \sigma^2 \Phi),$$

as Delgado and Mora (1994) prove. And, as a result from Theorem 1.a1, $H(\hat{\beta}_2)$ and $\hat{\Gamma}^{-1}$ converge in probability to $H(\alpha_2)$ and Γ^{-1}, respectively. Thus,

$$n^{1/2} \begin{pmatrix} \hat{\gamma}_1 - \beta_1 \\ \hat{\delta} \\ \hat{\gamma}_3 - \beta_3 \end{pmatrix} \xrightarrow{d} N(0, \sigma^2 \Gamma^{-1} H(\alpha_2)' \Phi H(\alpha_2) \Gamma^{-1}) = N(0, \sigma^2 \Gamma^{-1}). \qquad \blacksquare$$

Proof of Theorem 1.a3.- We have defined

$$t = n^{1/2}\hat{\delta}/(\hat{\sigma}^2\hat{a}_{(k+1)(k+1)})^{1/2}.$$

From Theorem 1.a2, we deduce that the numerator in t converges in distribution to a random variable with distribution $N(0,\sigma^2 a_{(k+1)(k+1)})$, where $a_{(k+1)(k+1)}$ is the $(k+1)$th diagonal element in Γ^{-1}. And from the definition of Γ and $\hat{\Gamma}$ we deduce that $\hat{a}_{(k+1)(k+1)}$ converges in probability to $a_{(k+1)(k+1)}$. It only remains to prove that $\hat{\sigma}^2$ converges in probability to σ^2. $\hat{\sigma}^2$ is an estimator obtained under H_C. Thus,

$$\hat{\sigma}^2 = n^{-1}\sum_i(\hat{\varepsilon}_{Yi}I_i - \hat{\varepsilon}'_{1i}\hat{\gamma}_1 I_i - \hat{\varepsilon}'_{2i}\hat{\beta}_2\hat{\delta}I_i - \hat{\varepsilon}'_{3i}\hat{\gamma}_3 I_i)^2 =$$

$$= n^{-1}\sum_i(\hat{\varepsilon}'_{1i}\beta_1 I_i + \hat{\varepsilon}'_{3i}\beta_3 I_i + \hat{\varepsilon}_{Ui}I_i - \hat{\varepsilon}'_{1i}\hat{\gamma}_1 I_i - \hat{\varepsilon}'_{2i}\hat{\beta}_2\hat{\delta}I_i - \hat{\varepsilon}'_{3i}\hat{\gamma}_3 I_i)^2 =$$

$$= n^{-1}\sum_i\{\hat{\varepsilon}'_{1i}(\beta_1-\hat{\gamma}_1)I_i + \hat{\varepsilon}'_{3i}(\beta_3-\hat{\gamma}_3)I_i + \hat{\varepsilon}_{Ui}I_i - \hat{\varepsilon}'_{2i}\hat{\beta}_2\hat{\delta}I_i\}^2.$$

In this expression all terms converge in probability to 0 (as a consequence of Theorem 1, a1 and a2) except $n^{-1}\sum_i\hat{\varepsilon}^2_{Ui}I_i$, which converges to σ^2 as is proved in Delgado and Mora (1994). ■

Proof of Theorem 1.b1.- This result is proved in Delgado and Mora (1994), because this is a pure partially linear model. ■

Proof of Theorem 1.b2.- As H_1 is true, $\hat{\varepsilon}_{Yi}I_i = \hat{\varepsilon}'_{2i}\beta_2 I_i + \hat{\varepsilon}'_{3i}\beta_3 I_i + \hat{\varepsilon}_{Ui}I_i =$

$$= (\hat{\varepsilon}'_{1i}\ \ \hat{\varepsilon}'_{2i}\hat{\beta}_2\ \ \hat{\varepsilon}'_{3i})\begin{pmatrix} 0 \\ 1 \\ \beta_3 \end{pmatrix} I_i + (\beta_2-\hat{\beta}_2)'\hat{\varepsilon}_{2i}I_i + \hat{\varepsilon}_{Ui}I_i.$$

Taking this expression into (A.6) we obtain

$$\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\delta} \\ \hat{\gamma}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \beta_3 \end{pmatrix} + \hat{\Gamma}^{-1}H(\hat{\beta}_2)'(n^{-1}\sum_i \begin{pmatrix} \hat{\varepsilon}_{1i} \\ \hat{\varepsilon}_{2i} \\ \hat{\varepsilon}_{3i} \end{pmatrix} \hat{\varepsilon}'_{2i}I_i)(\beta_2-\hat{\beta}_2) +$$

$$+\ \hat{\Gamma}^{-1}n^{-1}\textstyle\sum_1 \begin{pmatrix} \hat{\varepsilon}_{11} \\ \hat{\beta}_2'\hat{\varepsilon}_{21} \\ \hat{\varepsilon}_{31} \end{pmatrix} \hat{\varepsilon}_{U1} I_1 .$$

And in this expression the second and the third term converge to *0* in probability as a consequence of previous results. ∎

Proof of Theorem 1.b3.- As before,

$$t = n^{1/2}\hat{\delta}/(\hat{\sigma}^2\hat{a}_{(k+1)(k+1)})^{1/2}.$$

As in Theorem 1.a3, $\hat{\sigma}^2$ converges in probability to σ^2. Hence, as a consequence of Theorem 1.b2, $\hat{\delta}/(\hat{\sigma}^2\hat{a}_{(k+1)(k+1)})^{1/2}$ converges in probability to $(\sigma^2 a_{(k+1)(k+1)})^{-1/2}>0$. And Theorem 1.b3 follows straightforwardly from it. ∎

REFERENCES

Atkinson, A.B., Gomulka, J. and Stern, N.H. (1990), "Spending on Alcohol: Evidence from the Family Expenditure Survey 1970-1983", *Economic Journal,* 100, 808-827.

Banks, J., Blundell, R. and Lewbel, A. (1993), "Quadratic Engel Curves, Welfare Measurement and Consumer Demand", Manuscript.

Cox, D.R. (1961), "Tests of separate families of hypotheses" in *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability,* Vol.1, Univ. of Chicago Press, Berkeley, CA, 105-123.

Cox, D.R. (1962), "Further results on tests of separate families of hypotheses", *Journal of the Royal Statistical Society, B*, 24, 406-424.

Davidson, R. and Mackinnon, J. (1981), "Several tests for model specification in the presence of alternative hypotheses", *Econometrica*, 49, 781-793.

Deaton, A. (1981), "Three Essays on a Sri Lanka Household Survey", *Living Standard Measurement Study Working Paper no.11,* Washington: World Bank.

Deaton, A., Ruiz-Castillo, J. and Thomas, D. (1989), "The Influence of Household Composition on Household Expenditure Patterns: Theory and Spanish Evidence", *Journal of Political Economy*, 97, 179-200.

Delgado, M.A. and Stengos, T. (1994), "Semiparametric Specification Testing of Non-Nested Econometric Models", *Review of Economic Studies*, 207, 291-303.

Delgado, M.A. and Mora, J. (1994), "Nonparametric and Semiparametric Estimation with Discrete Regressors", Working Paper 94-11, Universidad Carlos III de Madrid.

Eicker, F. (1963), "Asymptotic normality and consistency of the least squares estimator for families of linear regressions", *Annals of Statistics*, 34, 447-456.

Fisher, G.R. (1983), "Tests for Two Separate Regressions", *Journal of Econometrics*, 21, 117-132.

Godfrey, L.G. and Pesaran, M.H. (1983), "Tests of Non-Nested Regression Models", *Journal of Econometrics,* 133-154.

Gozalo, P.L. (1992), "Nonparametric Analysis of Engel Curves: Estimation and Testing of Demographic Effects", Working Paper no. 92-15, Brown University.

Leser, C.E.V. (1963), "Forms of Engel Curve Functions", *Econometrica*, 31, 694-703.

Lewbel, A. (1991), "The Rank of Demand Systems: Theory and Nonparametric Estimation", *Econometrica*, 59, 377-391.

MacKinnon, J.G. (1990), "Model specification tests and artificial regression", *Journal of Economic Literature.*

MacKinnon, J.G., White, H. and Davidson, R. (1983), "Tests for Model Specification in the Presence of Alternative Hypothesis", *Journal of Econometrics.*

Nadaraya, E.A. (1964), "On estimating regression", *Theory of Probability and its Applications,* 9, 141-142.

Pesaran, M.H. (1982), "Comparison of local power of alternative tests of non-nested regression models", *Econometrica*, 50, 1287-1305.

Pesaran, M.H. and Deaton, A.S. (1978), "Testing non-nested nonlinear regression models", *Econometrica*, 46, 677-694.

Pollack, R.A. and Wales, T.J. (1978), "Estimation of Complete Demand Systems from Household Budged Data: The Linear and Quadratic Expenditure Systems", *American Economic Review,* 68, 348-359.

Pollack, R.A. and Wales, T.J. (1980), "Comparison of the Quadratic Expenditure System and Translog Demand Systems with Alternative Specifications of Demographic Effects, *Econometrica*, 48, 595-612.

Pollack, R.A. and Wales, T.J. (1981), "Demographic Variables in Demand Analysis", *Econometrica*, 49, 1533-1551.

Robinson, P.M. (1988), "Root-n-consistent semiparametric regression", *Econometrica*, 56, 931-954.

Speckman, P.(1988), "Kernel smoothing in partially linear models", *Journal of the Royal Statistical Society, B,* 50, 413-446.

Stone, C.J. (1977), "Consistent nonparametric regression", *Annals of Statistics*, 4, 595-645.

Watson, G.S. (1964), "Smooth regression analysis", *Sankhya A,* 26, 359-372.

White, H. (1980), "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity", *Econometrica*, 48, 817-838.

White, H. (1982), "Regularity conditions for Cox's test of non-nested hypotheses", *Journal of Econometrics*, 19, 301-318.

Working, H. (1943), "Statistical Laws of Family Expenditure", *Journal of the American Statistical Association*, 38, 43-56.

TABLE 1

Monte Carlo results for models 1-6 (Size of the test)

Sample Size = *40*, Number of replications = *10000*

		Empirical Level		$\hat{\gamma}_1$		$\hat{\gamma}_3$		$\hat{\delta}$	
		$\alpha=0.05$	$\alpha=0.01$	M	V	M	V	M	V
M1	Non-S	*0.0978*	*0.0346*	*4.0003* 4.0	*0.0694* 0.05	-	-	*-0.005* 0.0	*0.0096* 0.0063
	Kern. *h=1.2*	*0.1131*	*0.0369*	*4.1300* 4.0	*0.0697* 0.05	-	-	*0.053* 0.0	*0.0059* 0.0063
M2	Kern. *h=0.5*	*0.1043*	*0.0322*	*4.0165* 4.0	*0.0770* 0.05	-	-	*0.004* 0.0	*0.0102* 0.0063
	Kern. *h=0.8*	*0.0862*	*0.0286*	*4.0325* 4.0	*0.0745* 0.05	-	-	*0.013* 0.0	*0.0085* 0.0063
M3	Kern. *h=0.5*	*0.1030*	*0.0353*	*4.0213* 4.0	*0.0779* 0.05	-	-	*0.002* 0.0	*0.0102* 0.0063
	Kern. *h=0.8*	*0.0919*	*0.0297*	*4.0297* 4.0	*0.0734* 0.05	-	-	*0.015* 0.0	*0.0089* 0.0063
M4	Non-S	*0.0995*	*0.0311*	*4.0027* 4.0	*0.0748* 0.05	*1.0010* 1.0	*0.0362* 0.025	*-0.015* 0.0	*0.6890* 0.01
	Kern. *h=1.2*	*0.2840*	*0.1219*	*4.1123* 4.0	*0.0710* 0.05	*0.7771* 1.0	*0.0297* 0.025	*0.124* 0.0	*0.0075* 0.01
M5	Non-S	*0.0951*	*0.0340*	*4.0032* 4.0	*0.0706* 0.05	*-1.998* -2.0	*0.0357* 0.025	*-0.082*	*1099.7*
	Kern. *h=1.2*	*0.2799*	*0.1213*	*4.1122* 4.0	*0.0713* 0.05	*-2.221* -2.0	*0.0296* 0.025	*0.467*	*386.8*
M6	K./NS *h=10*	*0.1929*	*0.0973*	*0.2293* 0.229	*0.0160*	*-0.014* -0.014	*3.0E-5*	*0.001* 0.0	*0.0513*
	K./NS *h=60*	*0.0871*	*0.0294*	*0.2286* 0.229	*0.0079*	*-0.014* -0.014	*1.4E-5*	*-0.001* 0.0	*0.0237*

Notes.- 1. In small numbers the corresponding asymptotic values (See Section 2.3 and theorem 1)

2. In models 5 and 6 columns labelled $\hat{\gamma}_1$ and $\hat{\gamma}_3$ contain the results for the first and second coordinates of vector $\hat{\gamma}_1$.

TABLE 2

Monte Carlo results for models 1-6 (Size of the test)

Sample Size = *200*, Number of replications = *2000*

		Empirical Level $\alpha=0.05$	Empirical Level $\alpha=0.01$	$\hat{\gamma}_1$ M	$\hat{\gamma}_1$ V	$\hat{\gamma}_3$ M	$\hat{\gamma}_3$ V	$\hat{\delta}$ M	$\hat{\delta}$ V
M1	Non-S	*0.0590*	*0.0185*	*4.0015* 4.0	*0.0110* 0.01	-	-	*-0.001* 0.0	*0.0014* 0.0015
M1	Kern. *h=1.2*	*0.2845*	*0.1045*	*4.1152* 4.0	*0.0110* 0.01	-	-	*0.049* 0.0	*0.0010* 0.0015
M2	Kern. *h=0.4*	*0.0655*	*0.0160*	*4.0044* 4.0	*0.0110* 0.01	-	-	*0.002* 0.0	*0.0014* 0.0015
M2	Kern. *h=0.7*	*0.0570*	*0.0145*	*4.0122* 4.0	*0.0107* 0.01	-	-	*0.007* 0.0	*0.0013* 0.0015
M3	Kern. *h=0.4*	*0.0575*	*0.0135*	*4.0060* 4.0	*0.0109* 0.01	-	-	*0.001* 0.0	*0.0014* 0.0015
M3	Kern. *h=0.7*	*0.0655*	*0.0175*	*4.0109* 4.0	*0.0112* 0.01	-	-	*0.009* 0.0	*0.0013* 0.0015
M4	Non-S	*0.0615*	*0.0165*	*4.0039* 4.0	*0.0109* 0.01	*1.0006* 1.0	*0.0056* 0.005	*-0.004* 0.0	*0.0029* 0.002
M4	Kern. *h=1.2*	*0.8435*	*0.6340*	*4.0996* 4.0	*0.0115* 0.01	*0.7989* 1.0	*0.0453* 0.005	*0.117* 0.0	*0.0012* 0.002
M5	Non-S	*0.0590*	*0.0140*	*4.0034* 4.0	*0.0108* 0.01	*-1.998* -2.0	*0.0057* 0.005	*-0.718*	*2033.9*
M5	Kern. *h=1.2*	*0.8285*	*0.6160*	*4.0991* 4.0	*0.0109* 0.01	*-2.198* -2.0	*0.0048* 0.005	*0.638*	*0.3405*
M6	K./NS *h=10*	*0.0815*	*0.0285*	*0.2295* 0.229	*0.0011*	*-0.015* -0.015	*1.7E-6*	*0.002* 0.0	*0.0017*
M6	K./NS *h=60*	*0.0495*	*0.0105*	*0.2285* 0.229	*0.0008*	*-0.015* -0.015	*1.3E-6*	*0.0001* 0.0	*0.0014*

Notes.- 1. In small numbers the corresponding asymptotic values (See Section 2.3 and theorem 1)

2. In models 5 and 6 columns labelled $\hat{\gamma}_1$ and $\hat{\gamma}_3$ contain the results for the first and second coordinates of vector $\hat{\gamma}_1$.

TABLE 3

Monte Carlo results for models 7-12 (Power of the test)

Sample Size = *40*, Number of replications = *10000*

		Empirical Level		$\hat{\gamma}_1$		$\hat{\gamma}_3$		$\hat{\delta}$	
		$\alpha=0.05$	$\alpha=0.01$	M	V	M	V	M	V
M 7	Non-S	*0.6895*	*0.4964*	*0.0003* 0.0	*0.0694*	-	-	*1.0071* 1.00	*0.9958*
	Kern. *h=1.2*	*0.8229*	*0.6334*	*0.1300* 0.0	*0.0697*	-	-	*0.8850* 1.00	*0.0644*
M 8	Kern. *h=0.5*	*0.6834*	*0.5011*	*0.0165* 0.0	*0.0770*	-	-	*0.9710* 1.00	*0.5571*
	Kern. *h=0.8*	*0.7077*	*0.5086*	*0.0325* 0.0	*0.0746*	-	-	*0.9713* 1.00	*1.4938*
M 9	Kern. *h=0.5*	*0.9086*	*0.8206*	*0.0213* 0.0	*0.0779*	-	-	*0.9778* 1.00	*0.2285*
	Kern. *h=0.8*	*0.9269*	*0.8386*	*0.0297* 0.0	*0.0733*	-	-	*0.9769* 1.00	*0.0574*
M10	Non-S	*0.8890*	*0.7671*	*0.0002* 0.0	*0.0716*	*1.0017* 1.0	*0.0354*	*0.9889* 1.00	*0.8751*
	Kern. *h=1.2*	*0.9942*	*0.9734*	*0.1098* 0.0	*0.0682*	*0.7789* 1.0	*0.0295*	*0.9511* 1.00	*0.0169*
M11	Non-S	*0.9676*	*0.9152*	*0.0032* 0.0	*0.0706*	*0.0016* 0.00	*0.0357*	*0.9972* 1.0	*0.0577*
	Kern. *h=1.2*	*0.9996*	*0.9954*	*0.1122* 0.0	*0.0713*	*-0.222* 0.00	*0.0296*	*1.1165* 1.0	*0.0341*
M12	K./NS *h=10*	*0.9528*	*0.9192*	*-0.002* 0.029	*0.0593*	*6.1E-5* 0.0014	*1.1E-4*	*1.0007* 1.00	*0.0833*
	K./NS *h=60*	*0.9805*	*0.9586*	*0.0006* 0.029	*0.0260*	*-3E-05* 0.0014	*4.7E-5*	*0.9998* 1.00	*0.0368*

Notes.- 1. In small numbers the corresponding asymptotic values (See Section 2.3 and theorem 1)

2. In models 10 and 11 columns labelled $\hat{\gamma}_1$ and $\hat{\gamma}_3$ contain the results for the first and second coordinates of vector $\hat{\gamma}_1$.

<u>TABLE 4</u>

Monte Carlo results for models 7-12 (Power of the test)

Sample Size = *200*, Number of replications = *2000*

		Empirical Level		$\hat{\gamma}_1$		$\hat{\gamma}_3$		$\hat{\delta}$	
		$\alpha=0.05$	$\alpha=0.01$	M	V	M	V	M	V
M 7	Non-S	*0.9995*	*0.9960*	*0.0015* 0.0	*0.0110*	-	-	*0.9979* 1.00	*0.0182*
	Kern. *h=1.2*	*1.0000*	*0.9995*	*0.1152* 0.0	*0.0110*	-	-	*0.8945* 1.00	*0.0094*
M 8	Kern. *h=0.4*	*1.0000*	*0.9980*	*0.0044* 0.0	*0.0110*	-	-	*0.9943* 1.00	*0.0177*
	Kern. *h=0.7*	*1.0000*	*0.9990*	*0.0122* 0.0	*0.0107*	-	-	*0.9859* 1.00	*0.0157*
M 9	Kern. *h=0.4*	*1.0000*	*1.0000*	*0.0060* 0.0	*0.0109*	-	-	*0.9949* 1.00	*0.0076*
	Kern. *h=0.7*	*1.0000*	*1.0000*	*0.0109* 0.0	*0.0112*	- -	- -	*0.9909* 1.00	*0.0075*
M10	Non-S	*1.0000*	*1.0000*	*0.0039* 0.0	*0.0109*	*1.0007* 1.0	*0.0056*	*0.9976* 1.00	*0.0046*
	Kern. *h=1.2*	*1.0000*	*1.0000*	*0.0996* 0.0	*0.0115*	*0.7989* 1.0	*0.0049*	*0.9526* 1.00	*0.0027*
M11	Non-S	*1.0000*	*1.0000*	*0.0034* 0.0	*0.0108*	*0.0015* 0.00	*0.0057*	*0.9969* 1.0	*0.0084*
	Kern. *h=1.2*	*1.0000*	*1.0000*	*0.0992* 0.0	*0.0109*	*-0.198* 0.00	*0.0048*	*1.1050* 1.0	*0.0054*
M12	K./NS *h=10*	*1.0000*	*1.0000*	*-0.002* 0.029	*0.0017*	*1.0E-4* 0.0014	*3.2E-6*	*1.0024* 1.00	*0.0022*
	K./NS *h=60*	*1.0000*	*1.0000*	*-1E-04* 0.029	*0.0015*	*3.4E-6* 0.0014	*2.7E-6*	*1.0002* 1.00	*0.0019*

<u>Notes</u>.- 1. In small numbers the corresponding asymptotic values (See Section 2.3 and theorem 1)

2. In models 10 and 11 columns labelled $\hat{\gamma}_1$ and $\hat{\gamma}_3$ contain the results for the first and second coordinates of vector $\hat{\gamma}_1$.

TABLE 5

Variables	Moments and Quantiles						
	Mean	S.D.	5%	25%	50%	75%	95%
Total Exp.	865550	629550	185505	449784	731113	1112480	1977378
Share Food	0.3783	0.1656	0.1306	0.2583	0.3642	0.4847	0.6744
Age or R.P	50.5385	15.1074	27	38	50	62	76
Size of H.	3.6950	1.7918	1	2	4	5	7

TABLE 6

Semiparametric estimates for equations 1-6 (Working-Leser form)

		$\hat\beta_1$	SE($\hat\beta_1$)	HS($\hat\beta_1$)	$\hat\beta_2$	SE($\hat\beta_2$)	HS($\hat\beta_2$)
Eq. 1	(with constant)	5.0E-1	2.9E-2	4.2E-2	-2.3E-2	1.1E-3	1.6E-3
Eq. 2	Kernel h=10	5.4E-1	2.9E-2	4.4E-2	-2.5E-2	1.1E-3	1.6E-3
	Kernel h=60	5.2E-1	2.9E-2	4.2E-2	-2.4E-2	1.1E-3	1.6E-3
Eq. 3	Non-S.	2.3E-1	2.9E-2	5.0E-2	-1.5E-2	1.1E-3	1.8E-3
Eq. 4	Ker/NS h=10	3.1E-1	2.9E-2	5.1E-2	-1.8E-2	1.1E-3	1.9E-3
	Ker/NS h=60	2.6E-1	2.9E-2	5.0E-2	-1.6E-2	1.1E-3	1.8E-3
Eq. 5	Ker/NS h=10	3.6E-1	2.9E-2	5.1E-2	-1.9E-2	1.1E-3	1.9E-3
	Ker/NS h=60	3.2E-1	2.9E-2	5.0E-2	-1.8E-2	1.1E-3	1.9E-3
Eq. 6	Ker/NS h=10	3.5E-1	2.9E-2	5.0E-2	-1.9E-2	1.1E-3	1.9E-3
	Ker/NS h=60	3.1E-1	2.9E-2	4.9E-2	-1.7E-2	1.1E-3	1.8E-3

Note.- SE ≡ Standard Error; HS ≡ Heteroskedastic Consistent SE

TABLE 7

Semiparametric estimates for equations 7-12 (Q.E.S. form)

		$\hat\beta_1$	SE($\hat\beta_1$)	HS($\hat\beta_1$)	$\hat\beta_2$	SE($\hat\beta_2$)	HS($\hat\beta_2$)
Eq. 7	(with constant)	-1.2E-7	1.7E-9	3.5E-9	9.8E+3	4.8E+2	1.5E+3
Eq. 8	Kernel h=10	-1.2E-7	1.7E-9	3.4E-9	8.0E+3	4.9E+2	1.6E+3
	Kernel h=60	-1.2E-7	1.7E-9	3.4E-9	9.0E+3	4.8E+2	1.5E+3
Eq. 9	Non-S.	-1.3E-7	1.7E-9	4.2E-9	1.6E+4	5.0E+2	2.2E+3
Eq.10	Ker/NS h=10	-1.3E-7	1.6E-9	4.0E-9	1.3E+4	5.0E+2	2.1E+3
	Ker/NS h=60	-1.3E-7	1.7E-9	4.1E-9	1.5E+4	5.0E+2	2.2E+3
Eq.11	Ker/NS h=10	-1.2E-7	1.7E-9	3.9E-9	1.1E+4	5.0E+2	2.0E+3
	Ker/NS h=60	-1.2E-7	1.7E-9	4.0E-9	1.3E+4	5.0E+2	2.0E+3
Eq.12	Ker/NS h=10	-1.2E-7	1.6E-9	3.9E-9	1.2E+4	5.1E+2	2.0E+3
	Ker/NS h=60	-1.2E-7	1.7E-9	3.9E-9	1.4E+4	5.1E+2	2.0E+3

Note.- SE ≡ Standard Error; HS ≡ Heteroskedastic Consistent SE

TABLE 8

Results for Tests 1-12. Sample size = *23972* (all observations)

H_0: Working-Leser form *vs.* H_1: Q.E.S. form

		$\hat{\beta}_1$	SE($\hat{\beta}_1$)	$\hat{\beta}_2$	SE($\hat{\beta}_2$)	$\hat{\delta}$	SE($\hat{\delta}$)	HS($\hat{\delta}$)	
T.1		2.0E-1	1.4E-3	-1.2E-2	1.0E-4	-7.5E-2	5.9E-3	9.3E-3	*
T.2	K. *h=10*	8.9E-1	4.1E-2	-4.0E-2	1.7E-3	-5.3E-1	4.3E-2	5.1E-2	*
	K. *h=60*	8.4E-1	3.9E-2	-3.9E-2	1.6E-3	-5.2E-1	4.2E-2	5.0E-2	*
T.3	Non-S.	4.6E-1	3.2E-2	-2.6E-2	1.3E-3	-5.1E-1	3.1E-2	4.4E-2	*
T.4	K. *h=10*	5.7E-1	3.4E-2	-3.0E-2	1.4E-3	-4.9E-1	3.3E-2	4.6E-2	*
	K. *h=60*	5.0E-1	3.2E-2	-2.8E-2	1.3E-3	-5.1E-1	3.2E-2	4.5E-2	*
T.5	K. *h=10*	6.4E-1	3.5E-2	-3.2E-2	1.4E-3	-5.0E-1	3.6E-2	4.9E-2	*
	K. *h=60*	6.0E-1	3.4E-2	-3.1E-2	1.4E-3	-5.2E-1	3.4E-2	4.8E-2	*
T.6	K. *h=10*	5.9E-1	3.5E-2	-3.0E-2	1.4E-3	-4.7E-1	3.5E-2	5.0E-2	*
	K. *h=60*	5.5E-1	3.4E-2	-2.9E-2	1.4E-3	-4.9E-1	3.4E-2	4.9E-2	*

Note.- SE ≡ Standard Error; HS ≡ Heteroskedastic Consistent SE.
An asterisk means that the null hypothesis is rejected (α=0.05).

H_0: Q.E.S. form *vs.* H_1: Working-Leser form

		$\hat{\beta}_1$	SE($\hat{\beta}_1$)	$\hat{\beta}_2$	SE($\hat{\beta}_2$)	$\hat{\delta}$	SE($\hat{\delta}$)	HS($\hat{\delta}$)	
T.7		-8.E-9	1.2E-9	-7.0E+3	5.2E+2	1.1E+0	5.0E-3	6.6E-3	*
T.8	K. *h=10*	5.6E-8	4.9E-9	-5.8E+3	6.0E+2	1.5E+0	3.9E-2	4.2E-2	*
	K. *h=60*	5.6E-8	4.8E-9	-6.0E+3	6.1E+2	1.5E+0	3.9E-2	4.3E-2	*
T.9	Non-S.	5.8E-8	4.0E-9	-1.0E+4	6.9E+2	1.4E+0	2.8E-2	3.3E-2	*
T10	K. *h=10*	5.3E-8	4.1E-9	-8.5E+3	6.6E+2	1.4E+0	2.9E-2	3.5E-2	*
	K. *h=60*	5.7E-8	4.0E-9	-9.7E+3	6.8E+2	1.4E+0	2.8E-2	3.4E-2	*
T11	K. *h=10*	5.3E-8	4.2E-9	-7.8E+3	6.4E+2	1.4E+0	3.1E-2	3.7E-2	*
	K. *h=60*	5.6E-8	4.1E-9	-8.7E+3	6.6E+2	1.4E+0	3.0E-2	3.6E-2	*
T12	K. *h=10*	5.0E-8	4.2E-9	-7.5E+3	6.6E+2	1.4E+0	3.1E-2	3.9E-2	*
	K. *h=60*	5.3E-8	4.1E-9	-8.4E+3	6.8E+2	1.4E+0	3.0E-2	3.8E-2	*

Note.- SE ≡ Standard Error; HS ≡ Heteroskedastic Consistent SE.
An asterisk means that the null hypothesis is rejected (α=0.05).

TABLE 9

Results for Tests 1-12. Sample size = *6710* ("standard" households)

H_0: Working-Leser form *vs.* H_1: Q.E.S. form

		$\hat{\beta}_1$	SE($\hat{\beta}_1$)	$\hat{\beta}_2$	SE($\hat{\beta}_2$)	$\hat{\delta}$	SE($\hat{\delta}$)	HS($\hat{\delta}$)	
T.1		*2.1E-1*	*4.0E-3*	*-1.4E-2*	*2.7E-4*	*3.5E-3*	*1.6E-2*	*2.2E-2*	
T.2	K. *h=15*	*2.3E-1*	*7.9E-2*	*-1.8E-2*	*3.0E-3*	*-5.7E-1*	*9.1E-2*	*1.1E-1*	*
	K. *h=75*	*1.7E-1*	*8.0E-2*	*-1.6E-2*	*3.0E-3*	*-6.0E-1*	*9.2E-2*	*1.2E-1*	*
T.3	Non-S.	*1.3E-1*	*7.8E-2*	*-1.5E-2*	*3.0E-3*	*-6.1E-1*	*8.8E-2*	*1.1E-1*	*
T.4	K. *h=15*	*2.4E-1*	*7.8E-2*	*-1.8E-2*	*3.0E-3*	*-5.9E-1*	*8.8E-2*	*1.1E-1*	*
	K. *h=75*	*1.5E-1*	*7.8E-2*	*-1.5E-2*	*3.0E-3*	*-6.1E-1*	*8.8E-2*	*1.1E-1*	*
T.5	K. *h=15*	*2.9E-1*	*7.8E-2*	*-2.0E-2*	*3.0E-3*	*-5.9E-1*	*9.1E-2*	*1.1E-1*	*
	K. *h=75*	*2.5E-1*	*7.9E-2*	*-1.9E-2*	*3.0E-3*	*-6.1E-1*	*9.1E-2*	*1.2E-1*	*
T.6	K. *h=15*	*2.9E-1*	*7.8E-2*	*-2.0E-2*	*3.0E-3*	*-6.0E-1*	*9.1E-2*	*1.1E-1*	*
	K. *h=75*	*2.4E-1*	*7.9E-2*	*-1.8E-2*	*3.0E-3*	*-6.3E-1*	*9.1E-2*	*1.2E-1*	*

Note.- SE ≡ Standard Error; HS ≡ Heteroskedastic Consistent SE. An asterisk means that the null hypothesis is rejected (α=0.05).

H_0: Q.E.S. form *vs.* H_1: Working-Leser form

		$\hat{\beta}_1$	SE($\hat{\beta}_1$)	$\hat{\beta}_2$	SE($\hat{\beta}_2$)	$\hat{\delta}$	SE($\hat{\delta}$)	HS($\hat{\delta}$)	
T.7		*1.8E-9*	*2.0E-9*	*-2.5E+2*	*2.7E+3*	*1.0E+0*	*1.5E-2*	*2.0E-2*	*
T.8	K. *h=15*	*5.1E-8*	*8.6E-9*	*-2.6E+4*	*4.7E+3*	*1.5E+0*	*8.8E-2*	*1.1E-1*	*
	K. *h=75*	*5.2E-8*	*8.6E-9*	*-2.8E+4*	*4.8E+3*	*1.6E+0*	*8.9E-2*	*1.1E-1*	*
T.9	Non-S.	*5.4E-8*	*8.4E-9*	*-3.0E+4*	*4.8E+3*	*1.6E+0*	*8.5E-2*	*1.1E-1*	*
T10	K. *h=15*	*5.4E-8*	*8.4E-9*	*-2.8E+4*	*4.6E+3*	*1.6E+0*	*8.5E-2*	*1.0E-1*	*
	K. *h=75*	*5.5E-8*	*8.4E-9*	*-3.0E+4*	*4.8E+3*	*1.6E+0*	*8.5E-2*	*1.1E-1*	*
T11	K. *h=15*	*5.3E-8*	*8.5E-9*	*-2.6E+4*	*4.5E+3*	*1.6E+0*	*8.8E-2*	*1.1E-1*	*
	K. *h=75*	*5.4E-8*	*8.5E-9*	*-2.8E+4*	*4.6E+3*	*1.6E+0*	*8.8E-2*	*1.1E-1*	*
T12	K. *h=15*	*5.3E-8*	*8.4E-9*	*-2.7E+4*	*4.5E+3*	*1.6E+0*	*8.8E-2*	*1.1E-1*	*
	K. *h=75*	*5.5E-8*	*8.4E-9*	*-2.9E+4*	*4.6E+3*	*1.6E+0*	*8.8E-2*	*1.1E-1*	*

Note.- SE ≡ Standard Error; HS ≡ Heteroskedastic Consistent SE. An asterisk means that the null hypothesis is rejected (α=0.05).

TABLE 10

Results for Tests 1-12. Sample size = *16779*

(*(0.1,0.8)*-quantiles in expenditure)

H_0: Working-Leser form *vs.* H_1: Q.E.S. form

		$\hat\beta_1$	SE($\hat\beta_1$)	$\hat\beta_2$	SE($\hat\beta_2$)	$\hat\delta$	SE($\hat\delta$)	HS($\hat\delta$)
T.1		2.1E-1	6.7E-3	-1.4E-2	4.4E-4	-4.2E-2	3.1E-2	3.3E-2
T.2	K. h=10	6.2E-1	4.2E-1	-2.9E-2	1.9E-2	-9.4E-2	6.9E-1	7.2E-1
	K. h=60	5.0E-1	4.0E-1	-2.4E-2	1.8E-2	2.2E-2	7.1E-1	7.4E-1
T.3	Non-S.	-2.E-1	1.8E-1	-1.4E-3	6.0E-3	-1.5E-1	5.0E-1	5.3E-1
T.4	K. h=10	1.4E-1	1.7E-1	-1.2E-2	8.3E-3	-1.1E-1	5.3E-1	5.6E-1
	K. h=60	-9.E-2	1.6E-1	-4.3E-3	6.3E-3	-1.2E-1	5.1E-1	5.4E-1
T.5	K. h=10	2.3E-1	1.9E-1	-1.6E-2	9.4E-3	-1.5E-1	5.4E-1	5.8E-1
	K. h=60	6.1E-2	1.6E-1	-9.7E-3	7.4E-3	-1.4E-1	5.3E-1	5.6E-1
T.6	K. h=10	1.8E-1	1.8E-1	-1.3E-2	9.0E-3	-2.9E-2	5.4E-1	5.8E-1
	K. h=60	3.8E-2	1.6E-2	-8.2E-2	7.2E-3	-5.3E-2	5.3E-1	5.7E-1

Note.- SE ≡ Standard Error; HS ≡ Heteroskedastic Consistent SE. An asterisk means that the null hypothesis is rejected (α=0.05).

H_0: Q.E.S. form *vs.* H_1: Working-Leser form

		$\hat\beta_1$	SE($\hat\beta_1$)	$\hat\beta_2$	SE($\hat\beta_2$)	$\hat\delta$	SE($\hat\delta$)	HS($\hat\delta$)	
T.7		-8.E-9	6.4E-9	-6.3E+3	4.7E+3	1.0E+0	3.1E-2	3.2E-2	*
T.8	K. h=10	1.3E-8	1.1E-7	-1.9E+3	1.7E+4	1.1E+0	6.9E-1	7.2E-1	
	K. h=60	-3.E-9	1.1E-7	6.4E+2	1.8E+4	9.8E-1	7.1E-1	7.4E-1	
T.9	Non-S.	2.2E-8	7.6E-8	-7.4E+3	2.5E+4	1.1E+0	5.0E-1	5.3E-1	*
T10	K. h=10	1.8E-8	8.3E-8	-4.6E+3	2.2E+4	1.1E+0	5.3E-1	5.6E-1	*
	K. h=60	1.8E-8	7.8E-8	-5.6E+3	2.5E+4	1.1E+0	5.1E-1	5.4E-1	*
T11	K. h=10	2.3E-8	8.5E-8	-5.6E+3	2.1E+4	1.1E+0	5.4E-1	5.8E-1	
	K. h=60	2.1E-8	8.1E-8	-5.8E+3	2.3E+4	1.1E+0	5.3E-1	5.6E-1	*
T12	K. h=10	4.8E-9	8.4E-8	-1.1E+3	2.1E+4	1.0E+0	5.4E-1	5.8E-1	
	K. h=60	8.1E-9	8.0E-8	-2.2E+3	2.3E+4	1.1E+0	5.3E-1	5.5E-1	*

Note.- SE ≡ Standard Error; HS ≡ Heteroskedastic Consistent SE. An asterisk means that the null hypothesis is rejected (α=0.05).

Figure 1.- Density Function of Total Expenditure (Kernel Estimate)

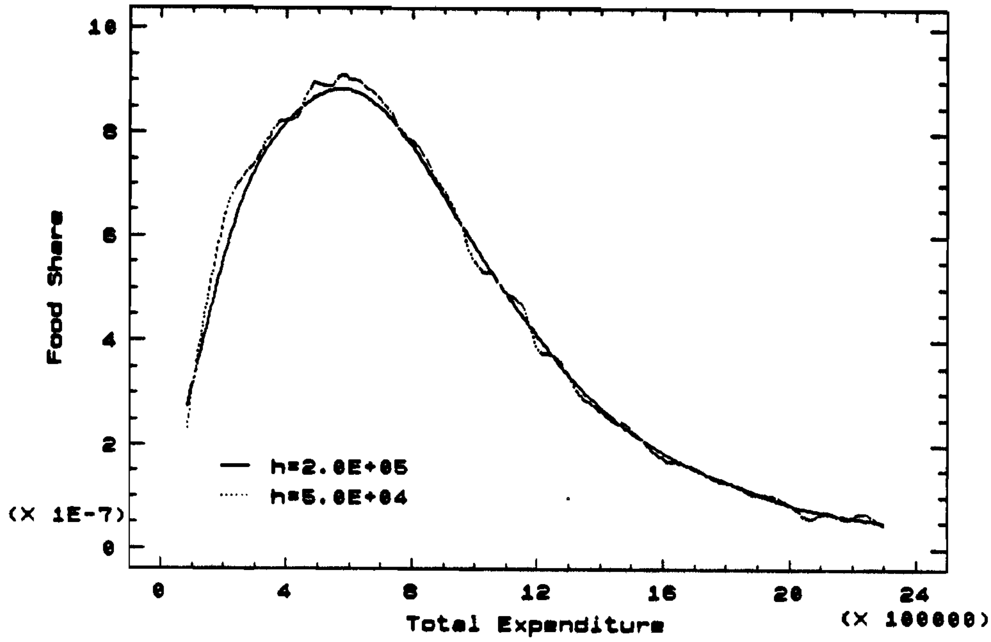

Figure 2.- Cross-Plot X/Y

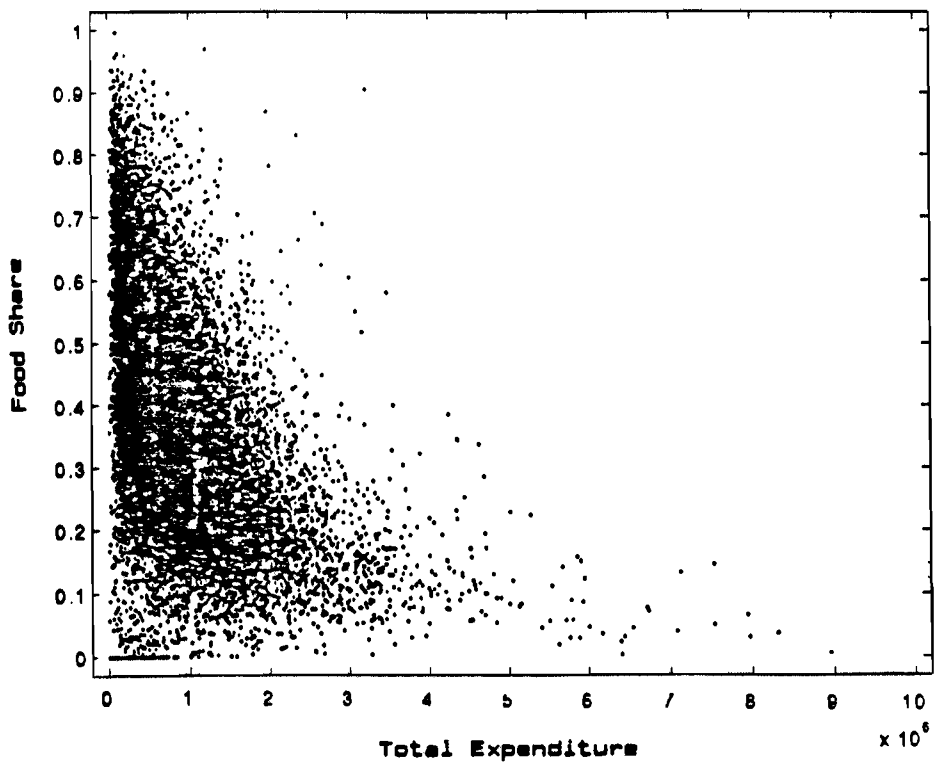

Figure 3a.- Food Share for All Households (Kernel Estimate)

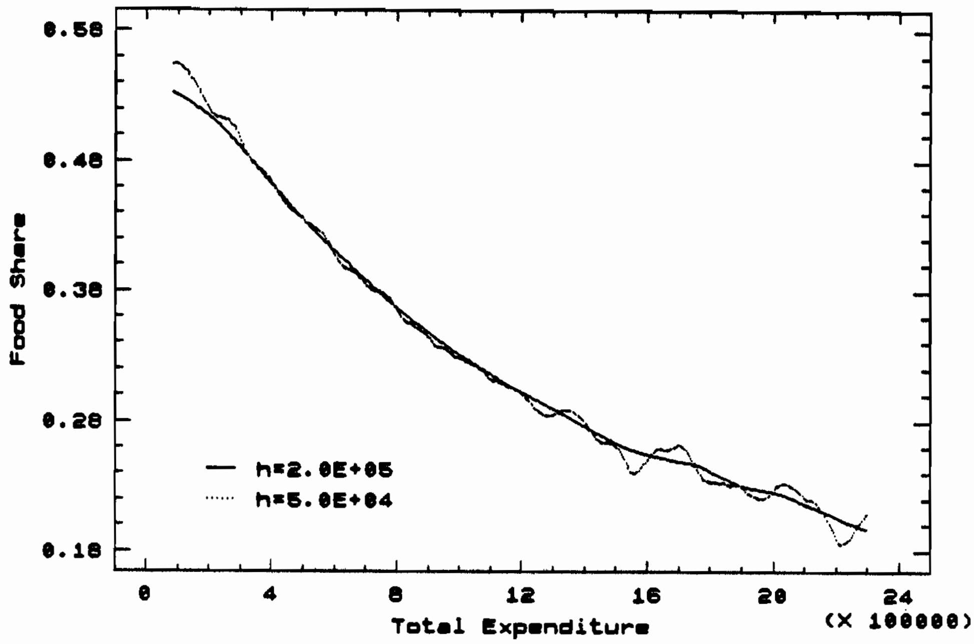

Figure 3b.- Food Share for All Households (*k*-NN Estimate)

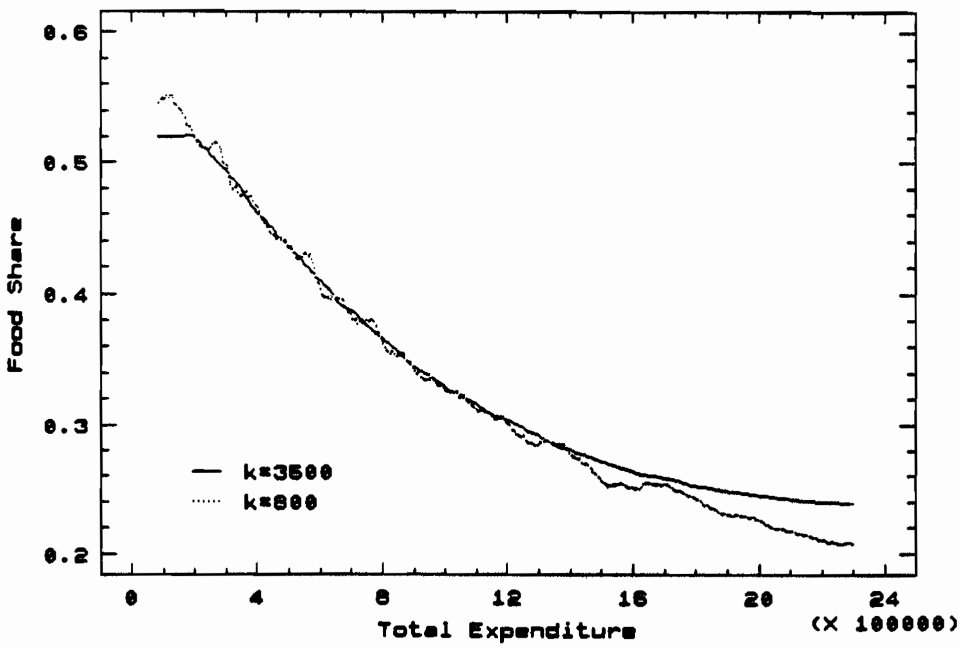

Figure 4.- Food Share by Household Type

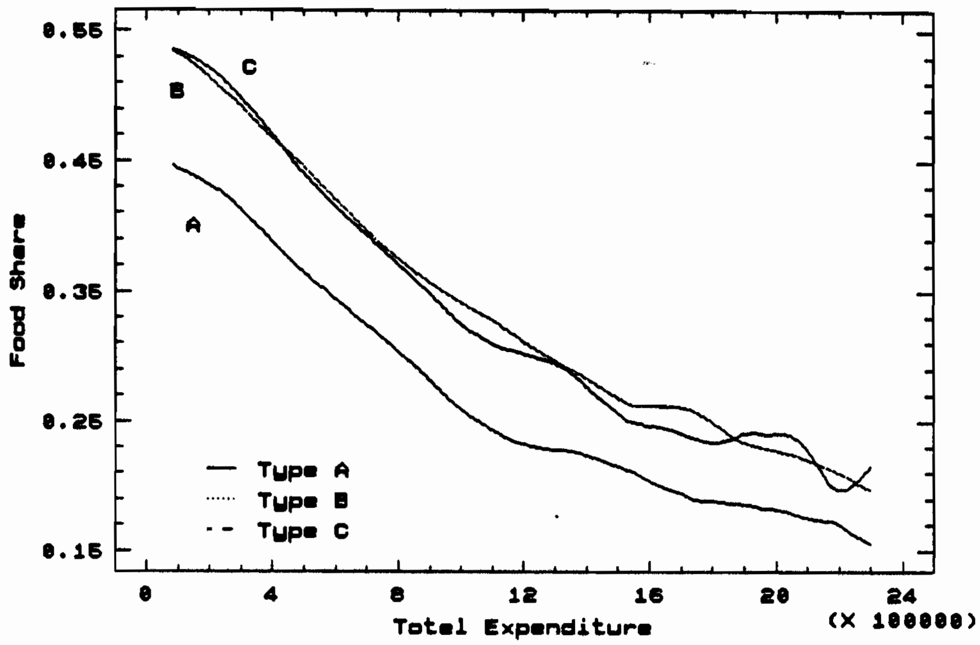

Figure 5.- Food Share by Household Size